

This Page Is Inserted by IFW Operations  
and is not a part of the Official Record

## **BEST AVAILABLE IMAGES**

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images may include (but are not limited to):

- BLACK BORDERS
- TEXT CUT OFF AT TOP, BOTTOM OR SIDES
- FADED TEXT
- ILLEGIBLE TEXT
- SKEWED/SLANTED IMAGES
- COLORED PHOTOS
- BLACK OR VERY BLACK AND WHITE DARK PHOTOS
- GRAY SCALE DOCUMENTS

**IMAGES ARE BEST AVAILABLE COPY.**

**As rescanning documents *will not* correct images,  
please do not report the images to the  
Image Problem Mailbox.**

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.3  
Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 9, 2002, 06:49:07 ; Search time 38 Seconds  
(without alignments)  
207.382 Million cell updates/sec

Title: US-09-895-298A-83  
Perfect score: 190  
Sequence: 1 MMNFQPSKAMRASQMTTF.....HDGSLDRSRVQEGNPRA 190

Scoring table: OLIGO  
Gapop 60.0 , Gapext 60.0

Searched: 112892 seqs, 41476328 residues

Word size : 4

Total number of hits satisfying chosen parameters: 49917

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Database : SwissProt\_40.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	7	3.7	115	1 SMS2_ONCMY	091194 oncorhynchu
2	7	3.7	136	1 YDP7_SCHPO	014010 schizosacch
3	7	3.7	151	1 SSB_RAT	P28042 rattus norv
4	7	3.7	215	1 Y042_METJA	060347 methanococc
5	7	3.7	241	1 YR89_METJA	058984 methanococc
6	7	3.7	259	1 IF2C_GALISU	008810 galdieria s
7	7	3.7	338	1 PHND_ECOLI	P16682 escherichia
8	7	3.7	500	1 YDGR_ECOLI	P77304 escherichia
9	7	3.7	550	1 SYM_CHLMU	09107 chlamydia m
10	7	3.7	553	1 ASNB_ECOLI	P22106 escherichia
11	7	3.7	1176	1 NTR_NEUCR	P38681 neurospora
12	7	3.7	1410	1 PDRB_YEAST	P40550 saccharomyc
13	6	3.2	20	1 YPRB_SERMA	P22581 serratia ma
14	6	3.2	61	1 HSP_CHICK	P15340 gallus gall
15	6	3.2	65	1 YC33_CIVAPA	P48273 cyanophora
16	6	3.2	67	1 PSBH_PORPU	P51325 porphyra pu
17	6	3.2	88	1 VG30_BPMD2	064224 mycobacteri
18	6	3.2	106	1 GAS4_ARATH	P46690 arabidopsis
19	6	3.2	112	1 KV2D_MOUSE	P01629 mus musculu
20	6	3.2	112	1 OL7G_MOUSE	Q60888 mus musculu
21	6	3.2	113	1 KVZE_MOUSE	P03976 mus musculu
22	6	3.2	113	1 RL30_SPOPR	P58375 spodoptera
23	6	3.2	114	1 RL30_BRABE	P58374 branchiosto
24	6	3.2	136	1 ACP1_CASGL	P93092 casuarina g
25	6	3.2	138	1 Y850_AQUAE	067017 aquilex aeo
26	6	3.2	139	1 Y4GB_RHISN	P55458 rhizobium s
27	6	3.2	146	1 AR16_CAEEL	P91167 caenorhabdi
28	6	3.2	148	1 SSB_HUMAN	004837 homo sapien
29	6	3.2	149	1 Y919_METJA	058329 methanococc
30	6	3.2	157	1 XY52_PSEPU	Q05092 pseudomonas
31	6	3.2	167	1 TETL_HUMAN	015273 homo sapien
32	6	3.2	167	1 TETL_MOUSE	070548 mus musculu
33	6	3.2	170	1 YC66_MESVI	Q9muq8 mesostigma

34	6	3.2	172	1 RS5_TREPA	083236 treponema p
35	6	3.2	177	1 SYC_BUCAP	P46241 buchnera ap
36	6	3.2	192	1 SIEB_BPP22	P38396 bacterioph
37	6	3.2	205	1 DHPS_CLOBE	Q05621 clostridium
38	6	3.2	205	1 RAB4_DICDI	P36410 dictyostell
39	6	3.2	207	1 KGUA_THEMEA	Q9x215 thermotoga
40	6	3.2	208	1 RH01_ENTHI	P31021 entamoeba h
41	6	3.2	215	1 CCMB_PARDE	P52219 paracoccus
42	6	3.2	215	1 RB14_HUMAN	P35287 homo sapien
43	6	3.2	218	1 CCMB_RHOCA	P29960 rhodobacter
44	6	3.2	220	1 Y008_METJA	Q60319 methanococc
45	6	3.2	221	1 HYPB_METJA	Q57884 methanococc

## ALIGNMENTS

RESULT 1					
SMS2_ONCMY		STANDARD:	PRT:	115 AA.	
ID SMS2_ONCMY					
AC 091194;					
DT 16-OCT-2001 (Rel. 40, Created)					
DT 16-OCT-2001 (Rel. 40, Last sequence update)					
DT 16-OCT-2001 (Rel. 40, Last annotation update)					
DE Somatostatin II precursor [Contains: [Tyr21,Gly24]somatostatin-28;					
DE [Tyr7,Gly10]somatostatin-14].					
OS Oncorhynchus mykiss (Rainbow trout) (Salmo gairdneri).					
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;					
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei;					
OC Protacanthopterygii; Salmoniformes; Salmonidae; Oncorhynchus.					
OX NCBI_TaxID=8022;					
RN [1]					
RP SEQUENCE FROM N.A.					
RX MEDLINE=95354921; PubMed=7628684;					
RA Moore C.A., Kittilson J.D., Dahl S.K., Sheridan M.A.;					
RT "Isolation and characterization of a cDNA encoding for					
RT preprosomatostatin containing [Tyr7, Gly10]-somatostatin-14 from the					
RT endocrine pancreas of rainbow trout, Oncorhynchus mykiss."					
RL Gen. Comp. Endocrinol. 98:253-261(1995).					
CC -1- FUNCTION: SOMATOSTATIN INHIBITS THE RELEASE OF SOMATOTROPIN.					
CC -1- SUBCELLULAR LOCATION: Secreted.					
CC -1- SIMILARITY: BELONGS TO THE SOMATOSTATIN FAMILY.					
CC					
CC This SWISS-PROT entry is copyright. It is produced through a collaboration					
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -					
CC the European Bioinformatics Institute. There are no restrictions on its					
CC use by non-profit institutions as long as its content is in no way					
CC modified and this statement is not removed. Usage by and for commercial					
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/					
CC or send an email to license@sib-sib.ch).					
CC					
DR EMBL; U32471; AAC59695.1; -					
DR InterPro; IPR004250; Somatostatin.					
DR pfam; PF03002; Somatostatin; 1.					
KW Cleavage on pair of basic residues; Hormone; Signal; Multigene family.					
FT SIGNAL	1	18			POTENTIAL.
FT PROPEP	19	87			POTENTIAL.
FT PEPTIDE	88	115			[TYR21, GLY24]SOMATOSTATIN-28 (POTENTIAL).
FT PEPTIDE	102	115			[TYR7, GLY10]SOMATOSTATIN-14.
FT DISULFID	104	115			BY SIMILARITY.
SQ SEQUENCE	115 AA;	12963 MW;	520595025FCA6D91 CRC64;		
Query Match	3.7%;	Score 7;	DB 1;	Length 115;	
Best Local Similarity	100.0%;	Pred. No. 4.1;			
Matches	7;	Conservative	0;	Mismatches	0;
Indels	0;	Gaps	0;		
Db	175 LDLSRR 181				
	30 LDLSRR 36				
RESULT 2					
YDP7_SCHPO					

0P501.A77

ID YDP7\_SCHPO STANDARD; PRT; 136 AA.  
AC 014010;  
DT 15-JUL-1998 (Rel. 36, Created)  
DT 15-JUL-1998 (Rel. 36, Last sequence update)  
DT 15-JUN-2002 (Rel. 41, Last annotation update)  
DE Hypothetical protein C29A4.07 in chromosome I.  
GN SPAC29A4.07.  
OS Schizosaccharomyces pombe (Fission yeast).  
OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;  
OC Schizosaccharomycetales; Schizosaccharomycetaceae;  
OC Schizosaccharomycetes.  
OX NCBI\_TaxID=4896;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=972;  
RX MEDLINE=21848401; PubMed=11859360;  
RA Wood V., Gwilliam R., Rajandream M.A., Lyne M., Lyne R., Stewart A.,  
RA Sgouros J., Peat N., Hayles J., Baker S., Basham D., Bowman S.,  
RA Brooks K., Brown D., Brown S., Chillingworth T., Churcher C.M.,  
RA Collins M., Connor R., Cronin A., Davis P., Feltwell T., Fraser A.,  
RA Gentles S., Goble A., Hamlin N., Harris D., Hidalgo J., Hodgson G.,  
RA Holroyd S., Hornsby T., Howarth S., Huckle E.J., Hunt S., Jagsels K.,  
RA James K., Jones L., Jones M., Leather S., McDonald S., McLean J.,  
RA Mooney P., Moule S., Mungall K., Murphy L., Niblett D., Odeli C.,  
RA Oliver K., O'Neill S., Pearson D., Quail M.A., Rabinowitsch E.,  
RA Rutherford K., Rutter S., Saunders D., Seeger K., Sharp S.,  
RA Skelton J., Simmonds M., Squares R., Stevens S., Stevens K.,  
RA Taylor K., Taylor R.G., Tivey A., Walsh S.V., Warren T., Whitehead S.,  
RA Woodward J., Volckaert E., Aert R., Robben J., Grymonprez B.,  
RA Weltjens I., Vanstreels G., Rieger M., Schaefer M., Mueller-Auer S.,  
RA Gabel C., Fuchs M., Fritz C., Holzer E., Moesti D., Hilbert H.,  
RA Borzym K., Langer I., Beck A., Lebrach H., Reinhardt R., Pohl T.M.,  
RA Eger P., Zimmermann W., Wedler H., Wandt R., Purrelle B.,  
RA Goffeau A., Cadieu E., Dreano S., Gloux S., Lelaure V., Mottier S.,  
RA Galibert F., Aves S.J., Xiang Z., Hunt C., Moore K., Hurst S.M.,  
RA Lucas M., Rochet M., Gaillardin C., Taliana V.A., Garzon A., Thode G.,  
RA Daga R.R., Cruzado L., Jimenez J., Sanchez M., del Rey F., Benito J.,  
RA Dominguez A., Revuelta J.L., Moreno S., Armstrong J., Forsburg S.L.,  
RA Cerutti L., Lowe T., McCombie W.R., Paulsen I., Potashkin J.,  
RA Shpakovski G.V., Ussery D., Barrell B.G., Nurse P.;  
RT "The genome sequence of Schizosaccharomyces pombe.";  
RL Nature 415:871-880(2002).  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
CC  
DR EMBL; 297210; CAB10134.1; -  
KW Hypothetical protein.  
SQ SEQUENCE 136 AA; 15382 MW; 52EA5A74C06C9FE1 CRC64;  
QY 125 INEGDKD 131  
DB 40 INEGDKD 46  
Query Match 3.7%; Score 7; DB 1; Length 136;  
Best Local Similarity 100.0%; Pred. No. 4.8;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 125 INEGDKD 131  
DB 40 INEGDKD 46  
RESULT 3  
SSB\_RAT STANDARD; PRT; 151 AA.  
AC P28042;  
DT 01-AUG-1992 (Rel. 23, Created)  
DT 01-OCT-1993 (Rel. 27, Last sequence update)  
DT 30-MAY-2000 (Rel. 39, Last annotation update)  
DE Single-stranded DNA-binding protein, mitochondrial precursor  
DE (Mt-SSB) (MtSSB) (P16).

109 b

GN SSBP1 OR SSBP.  
OS Rattus norvegicus (Rat).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.  
OX NCBI\_TaxID=10116;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Liver;  
RX MEDLINE=93246247; PubMed=8482537;  
RA Tiraniti V., Rocchi M., Didonato S., Zeviani M.;  
RT "Cloning of human and rat cDNAs encoding the mitochondrial single-  
RT stranded DNA-binding protein (SSB).";  
RL Gene 126:219-225(1993).  
RN [2]  
RP SEQUENCE FROM N.A.  
RC STRAIN=Sprague-Dawley;  
RX MEDLINE=98278843; PubMed=9611270;  
RA Gupta S., van Tuyle G.C.;  
RT "The gene and processed pseudogenes of the rat mitochondrial single-  
RT strand DNA-binding protein: structure and promoter strength  
RT analyses.";  
RL Gene 212:269-278(1998).  
RN [3]  
RP SEQUENCE OF 17-46.  
RX MEDLINE=91024186; PubMed=2221914;  
RA Hoke G.D., Pavco P.A., Ledwith B.J., van Tuyle G.C.;  
RT "Structural and functional studies of the rat mitochondrial single  
RT strand DNA binding protein p16.";  
RL Arch. Biochem. Biophys. 282:116-124(1990).  
CC -1- FUNCTION: THIS PROTEIN BINDS PREFERENTIALLY AND COOPERATIVELY  
CC TO SS-DNA. PROBABLY INVOLVED IN MITOCHONDRIAL DNA REPLICATION.  
CC -1- SUBUNIT: HOMOTETRAMER (BY SIMILARITY).  
CC -1- SUBCELLULAR LOCATION: Mitochondrial.  
CC -1- SIMILARITY: BELONGS TO THE SSB FAMILY.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
CC  
DR EMBL; M94557; AAA67315.1; -  
DR EMBL; AF043635; AAC18063.1; -  
DR EMBL; AF043630; AAC18063.1; JOINED.  
DR EMBL; AF043631; AAC18063.1; JOINED.  
DR EMBL; AF043632; AAC18063.1; JOINED.  
DR EMBL; AF043633; AAC18063.1; JOINED.  
DR EMBL; AF043634; AAC18063.1; JOINED.  
DR PIR; S13264; S13264.  
DR PIR; JN0569; JN0569.  
DR HSP; Q04837; 3UUL.  
DR InterPro; IPR000424; SSB\_protein.  
DR Pfam; PF00436; SSB.1.  
DR TIGRfams; TIGR00621; ssb.1.  
DR PROSITE; PS00735; SSB\_1; 1.  
DR PROSITE; PS00736; SSB\_2; 1.  
KW DNA-binding; DNA replication; Mitochondrion; Transit peptide.  
FT TRANSIT 1  
FT CHAIN 17 151  
SQ SEQUENCE 151 AA; 17455 MW; B9CA830AB280D6D9 CRC64;  
QY 151 SSVLER 157  
DB 22 SSVLER 28  
Query Match 3.7%; Score 7; DB 1; Length 151;  
Best Local Similarity 100.0%; Pred. No. 5.3;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 151 SSVLER 157  
DB 22 SSVLER 28  
RESULT 4

```
Y042_METJA
ID Y042_METJA STANDARD; PRT; 215 AA.
AC Q60347;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Hypothetical protein MJ0042.
GN MJ0042.
OS Methanococcus jannaschii.
OC Archaea; Euryarchaeota; Methanococci; Methanococcales;
OC Methanocaldococcaceae; Methanocaldococcus.
OX NCBI_TaxID=2190;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-JAL-1 / DSM 2661 / ATCC 43067;
RX MEDLINE=96337999; PubMed=8688087;
RA Bult C.J., White O., Olsen G.J., Zhou L., Fleischmann R.D.,
RA Sutton G.G., Blake O., Fitzgerald L.M., Clayton R.A., Gocayne J.D.,
RA Kerlavage A.R., Dougherty B.A., Tomb J.-F., Adams M.D., Reich C.I.,
RA Overbeek R., Kirkness E.F., Weinstock K.G., Merrick J.M., Glodek A.,
RA Scott J.L., Geoghagen N.S.M., Weidman J.F., Sadow P.W., Hanna M.C.,
RA Uterback T.R., Kelley J.M., Peterson J.D., Sadow P.W., Nguyen D.,
RA Cotton M.D., Roberts K.M., Hurst M.A., Kaine B.P., Borodovsky M.,
RA Klenk H.-P., Fraser C.M., Smith H.O., Woese C.R., Venter J.C.;
RT "Complete genome sequence of the methanogenic archaeon, Methanococcus
RT jannaschii."
RL Science 273:1058-1073(1996).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: U67462; AAB98028.1; -.
CC TIGR: MJ0042; -.
DR Hypothetical protein; Complete proteome.
KW SEQUENCE 215 AA; 24892 MW; D1F39FD383A3A0A1 CRC64;
SQ
Query Match 3.7%; Score 7; DB 1; Length 215;
Best Local Similarity 100.0%; Pred. No. 7.3;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 134 LIEKLIK 140
Db 198 LIEKLIK 204
RESULT 5
YF89_METJA
ID YF89_METJA STANDARD; PRT; 241 AA.
AC Q58984;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Hypothetical protein MJ1589.
GN MJ1589.
OS Methanococcus jannaschii.
OC Archaea; Euryarchaeota; Methanococci; Methanococcales;
OC Methanocaldococcaceae; Methanocaldococcus.
OX NCBI_TaxID=2190;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-JAL-1 / DSM 2661 / ATCC 43067;
RX MEDLINE=96337999; PubMed=8688087;
RA Bult C.J., White O., Olsen G.J., Zhou L., Fleischmann R.D.,
RA Sutton G.G., Blake J.A., Fitzgerald L.M., Clayton R.A., Gocayne J.D.,
RA Kerlavage A.R., Dougherty B.A., Tomb J.-F., Adams M.D., Reich C.I.,
RA Overbeek R., Kirkness E.F., Weinstock K.G., Merrick J.M., Glodek A.,
RA Scott J.L., Geoghagen N.S.M., Weidman J.F., Fuhrmann J.L., Nguyen D.,
RA Uterback T.R., Kelley J.M., Peterson J.D., Sadow P.W., Hanna M.C.,
```

```
RA Cotton M.D., Roberts K.M., Hurst M.A., Kaine B.P., Borodovsky M.,
RA Klenk H.-P., Fraser C.M., Smith H.O., Woese C.R., Venter J.C.;
RT "Complete genome sequence of the methanogenic archaeon, Methanococcus
RT jannaschii."
RL Science 273:1058-1073(1996).
CC -1- SUBCELLULAR LOCATION: Integral membrane protein (Potential).
CC -1- SIMILARITY: TO MJ0871, MJ0880 AND MJ1556.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: U67599; AAB99617.1; -.
CC TIGR: MJ1589; -.
DR Hypothetical protein; Transmembrane; Complete proteome.
KW TRANSMEM 1 21 POTENTIAL.
FT TRANSMEM 43 63 POTENTIAL.
FT TRANSMEM 75 95 POTENTIAL.
FT TRANSMEM 108 128 POTENTIAL.
FT TRANSMEM 160 180 POTENTIAL.
FT TRANSMEM 200 220 POTENTIAL.
SQ SEQUENCE 241 AA; 27082 MW; 45C5ARD0496723EF CRC64;
Query Match 3.7%; Score 7; DB 1; Length 241;
Best Local Similarity 100.0%; Pred. No. 8.2;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 99 VLITTYL 105
Db 168 VLITTYL 174
RESULT 6
IF2C_GALSU
ID IF2C_GALSU STANDARD; PRT; 259 AA.
AC Q08810;
DT 01-OCT-1994 (Rel. 30, Created)
DT 01-OCT-1994 (Rel. 30, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Translation initiation factor IF-2, chloroplast (Fragment).
GN INFB.
OS Galdieria sulphuraria.
OC Chloroplast.
OC Eukaryota; Rhodophyta; Bangiophyceae; Porphyridiales; Porphyridiaceae;
OC Galdieria.
OX NCBI_TaxID=130081;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-14-1-1 / Isolate 107.79/Goettlingen;
RX MEDLINE=94033298; PubMed=8219057;
RA Kustrzewa M., Zetsche K.;
RT "Organization of plastid-encoded ATPase genes and flanking regions
RT including homologues of infb and tsf in the thermophilic red alga
RT Galdieria sulphuraria."
RL Plant Mol. Biol. 23:67-76(1993).
CC -1- FUNCTION: IF-2, ONE OF THE ESSENTIAL COMPONENTS FOR THE INITIATION
CC OF PROTEIN SYNTHESIS IN VITRO, PROTECTS FORMYL METHIONYL-TRNA FROM
CC SPONTANEOUS HYDROLYSIS AND PROMOTES ITS BINDING TO THE 30S
CC RIBOSOMAL SUBUNITS. IT IS ALSO INVOLVED IN THE HYDROLYSIS OF GTP
CC DURING THE FORMATION OF THE 70S RIBOSOMAL COMPLEX (BY SIMILARITY).
CC -1- SUBCELLULAR LOCATION: Chloroplast.
CC -1- SIMILARITY: BELONGS TO THE IF-2 FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
```

```
CC or send an email to license@sib.ch).
CC -----
CC DR EMBL; X66698; CAA47239.1; -.
CC DR PIR; S36410; S36410.
CC DR InterPro; IPR000795; EF_GTPbind.
CC DR InterPro; IPR000178; IF2.
CC DR Pfam; PF00009; GTP_EFTU; 1.
CC DR PROSITE; PS01176; IF2; PARTIAL.
CC FT Initiation factor; Protein biosynthesis; GTP-binding; Chloroplast.
CC FT DOMAIN 172 >259 G-DOMAIN.
CC FT NP_BIND 180 187 GTP (BY SIMILARITY).
CC FT NON_TER 259 259
CC SQ SEQUENCE 259 AA; 29746 MW; 803DC24390F42EBA CRC64;

Query Match
Best Local Similarity 3.7%; Score 7; DB 1; Length 259;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 134 LIEKLIK 140
Db 189 LIEKLIK 195

RESULT 7
PHND_ECOLI
ID PHND_ECOLI STANDARD; PRT; 338 AA.
AC P16682;
DT 01-AUG-1990 (Rel. 15, Created)
DT 01-NOV-1991 (Rel. 20, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Phosphonates-binding periplasmic protein precursor.
GN PHND OR PSID OR B4105.
OS Escherichia coli.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Escherichia.
OX NCBI_TaxID=562;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=K12;
RX MEDLINE=91193228; PubMed=1840580;
RA Makino K., Kim S.K., Shinagawa H., Amemura M., Nakata A.;
RT "Molecular analysis of the cryptic and functional phn operons for
RL phosphonate use in Escherichia coli K-12.";
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=K12 / MG1655;
RX MEDLINE=95334362; PubMed=7610040;
RA Burland V.D., Plunkett G. III, Sofia H.J., Daniels D.L.,
RA Blattner F.R.;
RT "Analysis of the Escherichia coli genome VI: DNA sequence of the
RL region from 92.8 through 100 minutes.";
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=B;
RX MEDLINE=90170953; PubMed=2155230;
RA Chen C.-M., Ye Q.-Z., Zhu Z., Wanner B.L., Walsh C.T.;
RT "Molecular biology of carbon-phosphorus bond cleavage. Cloning and
RT uptake and C-P lyase activity in Escherichia coli B.";
RL J. Biol. Chem. 265:4461-4471(1990).
CC -1- FUNCTION: PART OF THE BINDING-PROTEIN-DEPENDENT TRANSPORT SYSTEM
CC FOR ALKYLPHOSPHONATES.
CC -1- SUBCELLULAR LOCATION: Periplasmic (Potential).
CC -1- MISCELLANEOUS: THE SEQUENCE SHOWN IS THAT OF STRAIN K12.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/)
```

```
CC or send an email to license@sib.ch).
CC -----
CC DR EMBL; D90227; BAAL4263.1; -.
CC DR EMBL; U14003; AAA97004.1; -.
CC DR EMBL; AE000482; AAC77066.1; -.
CC DR EMBL; J05260; AAA24340.1; -.
CC DR PIR; E35718; E35718.
CC DR Ecogene; EG10714; phnd.
CC DR TIGRFAMS; TIGR01098; 3A0109s03r; 1.
CC KW Alkylphosphonate uptake; Transport; Periplasmic; Signal;
CC Complete proteome.
CC FT SIGNAL 1 26 POTENTIAL.
CC FT CHAIN 27 338 PHOSPHONATES-BINDING PERIPLASMIC PROTEIN.
CC FT VARIANT 312 312 E -> A (IN STRAIN B).
CC SQ SEQUENCE 338 AA; 37370 MW; 84B4366AE8D1BF62 CRC64;

Query Match
Best Local Similarity 3.7%; Score 7; DB 1; Length 338;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 141 LQDMEKK 147
Db 54 LQDMEKK 60

RESULT 8
YDGR_ECOLI
ID YDGR_ECOLI STANDARD; PRT; 500 AA.
AC P77304;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Hypothetical transporter ydgr.
GN YDGR OR B1634.
OS Escherichia coli.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Escherichia.
OX NCBI_TaxID=562;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=K12 / MG1655;
RX MEDLINE=97426617; PubMed=9278503;
RA Blattner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,
RA Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,
RA Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,
RA Mau B., Shao Y.;
RT "The complete genome sequence of Escherichia coli K-12.";
RL Science 277:1453-1474(1997).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=K12;
RX MEDLINE=97251357; PubMed=9097039;
RA Alba H., Baba T., Fujita K., Hayashi K., Inada T., Isono K.,
RA Itoh T., Kasai H., Kashimoto K., Kimura S., Kitakawa M.,
RA Kitagawa M., Makino K., Miki T., Mizobuchi K., Mori H., Mori T.,
RA Motomura K., Nakade S., Nakamura Y., Nashimoto H., Nishio Y.,
RA Oshima T., Saito N., Sampei G., Seki Y., Sivasubram S.,
RA Tagami H., Takeda J., Takemoto K., Takeuchi Y., Wada C.,
RA Yamamoto Y., Horiuchi T.;
RT "A 570-kb DNA sequence of the Escherichia coli K-12 genome
RT corresponding to the 28.0-40.1 min region on the linkage map.";
RL DNA Res. 3:363-377(1996).
CC -1- SUBCELLULAR LOCATION: Integral membrane protein (Potential).
CC -1- SIMILARITY: BELONGS TO THE PTR2 FAMILY OF TRANSPORTERS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/)
CC or send an email to license@sib.ch).
```

DR EMBL: AE002259; AAC74706.1; .  
DR EMBL: D90807; BAA15395.1; .  
DR EcoGene; EG13939; ydgr.  
DR InterPro; IPR000109; PTR2.  
DR InterPro; IPR005279; PEPH\_symporter.  
DR Pfam; PF00854; PTR2; 1.  
DR TIGRfams; TIGR00924; yjdl\_subl\_fam; 1.  
DR PROSITE; PS01022; PTR2\_1; 1.  
DR PROSITE; PS01023; PTR2\_2; 1.  
KW Hypothetical protein; Transpos; Transmembrane; Complete proteome.  
FT TRANSMEM 35 55 POTENTIAL.  
FT TRANSMEM 60 80 POTENTIAL.  
FT TRANSMEM 90 110 POTENTIAL.  
FT TRANSMEM 112 132 POTENTIAL.  
FT TRANSMEM 154 174 POTENTIAL.  
FT TRANSMEM 179 199 POTENTIAL.  
FT TRANSMEM 221 241 POTENTIAL.  
FT TRANSMEM 247 267 POTENTIAL.  
FT TRANSMEM 275 295 POTENTIAL.  
FT TRANSMEM 321 341 POTENTIAL.  
FT TRANSMEM 353 373 POTENTIAL.  
FT TRANSMEM 379 399 POTENTIAL.  
FT TRANSMEM 415 435 POTENTIAL.  
FT TRANSMEM 460 480 POTENTIAL.  
SQ SEQUENCE 500 AA; 53991 MW; BFDCE0BFCC206C46 CRC64;  
  
Query Match 3.7%; Score 7; DB 1; Length 500;  
Best local Similarity 100.0%; Pred. No. 16;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 147 KANPSSL 153  
Db 130 KANPSSL 136  
|||||  
  
RESULT 9  
SYM\_CHLMU STANDARD; PRT; 550 AA.  
ID SYM\_CHLMU  
AC Q9PL07;  
DT 16-OCT-2001 (Rel. 40, Created)  
DT 16-OCT-2001 (Rel. 40, Last sequence update)  
DT 16-OCT-2001 (Rel. 40, Last annotation update)  
DE Methionyl-tRNA synthetase (EC 6.1.1.10) (Methionine--tRNA ligase)  
DE (MetRS).  
GN METG OR TC0301.  
OS Chlamydia muridarum.  
OC Bacteria; Chlamydiales; Chlamydiaceae; Chlamydia.  
OX NCBI\_TaxID=83560;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=Mopn / N199;  
RX MEDLINE=20150255; PubMed=10684935;  
RA Read T.D., Brunham R.C., Shen C., Gill S.R., Heidelberg J.F.,  
RA White O., Hickey E.K., Peterson J., Utterback T., Berry K., Bass S.,  
RA Linher K., Weidman J., Khouri H., Craven B., Bowman C., Dodson R.,  
RA Gwin M., Nelson W., Deboy R., Kolonay J., McClarty G., Salzberg S.L.,  
RA Eisen J., Fraser C.M.;  
RT "Genome sequences of Chlamydia trachomatis Mopn and Chlamydia  
RT pneumoniae AR39.";  
RL Nucleic Acids Res. 28:1397-1406(2000).  
CC -I- FUNCTION: IT IS PROBABLY ESSENTIAL FOR CELL SURVIVAL, BEING  
CC REQUIRED NOT ONLY FOR ELONGATION OF PROTEIN SYNTHESIS BUT ALSO  
CC FOR THE INITIATION OF ALL MRNA TRANSLATION THROUGH INITIATOR  
CC tRNA(FMET) AMINOACYLATION.  
CC -I- CATALYTIC ACTIVITY: ATP + L-methionine + tRNA(Met) = AMP +  
CC diphosphate + L-methionyl-tRNA(Met).  
CC -I- COFACTOR: BINDS 1 ZINC ION (BY SIMILARITY).  
CC -I- SUBUNIT: HOMODIMER (BY SIMILARITY).  
CC -I- SUBCELLULAR LOCATION: Cytoplasmic.  
CC -I- SIMILARITY: BELONGS TO CLASS-I AMINOACYL-TRNA SYNTHETASE FAMILY.  
CC STRONG, TO CYSTEINYL-TRNA SYNTHETASE.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL: AE002297; AAF39166.1; .  
DR HSSP; P00959; 1MEA.  
DR TIGR; TC0301; .  
DR InterPro; IPR002300; tRNA-synt\_1a.  
DR InterPro; IPR001412; tRNA-synt\_1.  
DR InterPro; IPR002304; tRNA-synt\_met.  
DR Pfam; PF00133; tRNA-synt\_1; 1.  
DR PRINTS; PRO1041; TRNASYNTHMET.  
DR TIGRfams; TIGR00398; metc; 1.  
DR PROSITE; PS00178; AA\_TRNA\_LIGASE\_I; FALSE\_NEG.  
KW Aminoacyl-tRNA synthetase; Protein biosynthesis; Ligase; ATP-binding;  
KW Metal-binding; Zinc; Complete proteome.  
FT SITE 13 23 "HIGH" REGION.  
FT SITE 331 335 "KMSKS" REGION.  
FT BINDING 334 334 ATP (BY SIMILARITY).  
FT METAL 145 145 ZINC (BY SIMILARITY).  
FT METAL 148 148 ZINC (BY SIMILARITY).  
FT METAL 158 158 ZINC (BY SIMILARITY).  
FT METAL 161 161 ZINC (BY SIMILARITY).  
SQ SEQUENCE 550 AA; 62814 MW; 93D0DE14CF23C53C CRC64;  
  
Query Match 3.7%; Score 7; DB 1; Length 550;  
Best local Similarity 100.0%; Pred. No. 17;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 173 GSLDLRS 179  
Db 509 GSLDLRS 515  
|||||  
  
RESULT 10  
ASNB\_ECOLI STANDARD; PRT; 553 AA.  
ID ASNB\_ECOLI  
AC P22106;  
DT 01-AUG-1991 (Rel. 19, Created)  
DT 30-MAY-2000 (Rel. 39, Last sequence update)  
DT 15-JUN-2002 (Rel. 41, Last annotation update)  
DE Asparagine synthetase B [glutamine-hydrolyzing] (EC 6.3.5.4).  
DE ASNB OR B0674.  
GN ASNB\_ECOLI.  
OS Escherichia coli.  
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;  
OC Escherichia.  
OX NCBI\_TaxID=562;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=K12;  
RX MEDLINE=90330624; PubMed=1973930;  
RA Scofield M.A., Lewis W.S., Schuster S.M.;  
RT "Nucleotide sequence of Escherichia coli asnb and deduced amino acid  
RT sequence of asparagine synthetase B.";  
RL J. Biol. Chem. 265:12895-12902(1990).  
RN [2]  
RP SEQUENCE FROM N.A.  
RC STRAIN=K12 / MG1655;  
RX MEDLINE=97426617; PubMed=9278503;  
RA Blatter F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,  
RA Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,  
RA Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,  
RA Mau B., Shao Y.;  
RT "The complete genome sequence of Escherichia coli K-12.";  
RT Science 277:1453-1474(1997).  
RN [3]  
RP SEQUENCE FROM N.A.  
RC STRAIN=K12;  
RX MEDLINE=97061202; PubMed=8905232;  
RA Oshima T., Aiba H., Baba T., Fujita K., Hayashi K., Honjo A.,

```

RA Ikemoto K., Inada T., Itoh T., Kajihara M., Kanai K., Kashimoto K.,
RA Kimura S., Kitagawa M., Makino K., Masuda S., Miki T., Mizobuchi K.,
RA Mori H., Motomura K., Nakamura Y., Nishimoto H., Nishio Y., Saito N.,
RA Sempel G., Seki Y., Tagami H., Takemoto K., Wada C., Yamamoto Y.,
RA Yano M., Horuchi T.;
RT "A 718-kb DNA sequence of the Escherichia coli K-12 genome
RT corresponding to the 12,728.0 min region on the linkage map.";
RL DNA Res. 3:137-155(1996).
RN [4]
RP X-RAY CRYSTALLOGRAPHY (2.0 ANGSTROMS).
RX MEDLINE-20056034; PubMed-10587437;
RA Larsen T.M., Boehlein S.K., Schuster S.M., Richards N.G.J.,
RA Thoden J.B., Holden H.M., Rayment I.;
RT "Three-dimensional structure of Escherichia coli asparagine
RT synthetase B: a short journey from substrate to product.";
RL Biochemistry 38:16146-16157(1999).
CC -1- CATALYTIC ACTIVITY: ATP + L-aspartate + L-glutamine -> AMP +
CC diphosphate + L-asparagine + L-glutamate.
CC -1- PATHWAY: Asparagine biosynthesis.
CC -1- SUBUNIT: HOMODIMER.
CC -1- MISCELLANEOUS: THIS ENZYME CAN USE EITHER AMMONIA OR GLUTAMINE AS
CC A SUBSTRATE WITH GLUTAMINE BEING THE PREFERRED NITROGEN SOURCE.
CC -1- SIMILARITY: CONTAINS 1 TYPE-2 GLUTAMINE AMIDOTRANSFERASE DOMAIN.
CC -1- SIMILARITY: TO OTHER ASN SYNTHETASES.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; J05554; AAA23498.1; -.
DR EMBL; AE000171; AAC73768.1; -.
DR EMBL; D90706; BAA35317.1; -.
DR PIR; A36616; AJECN.
DR PDB; 1CT9; 15-DEC-99.
DR EcoGene; EGI0092; asnB.
DR InterPro; IPR001962; Asn_synthase.
DR InterPro; IPR000583; GATase_2.
DR Pfam; PF00310; GATase_2; 1.
DR Pfam; PF00733; Asn_synthase; 1.
DR PROSITE; PS00443; GATASE_TYPE_II; 1.
DR PROSITE; PS00443; GATASE_TYPE_II; 1.
KW Ligase; Asparagine biosynthesis; Glutamine amidotransferase;
KW 3D-structure; Complete proteome.
FT INIT_MER 0
FT DOMAIN 1 174 GLUTAMINE AMIDOTRANSFERASE.
FT ACT_SITE 1 1 GATASE.
FT SEQUENCE 553 AA; 62527 MW; 908BD02A23EC565C CRC64;
SQ
Query Match 3.7%; Score 7; DB 1; Length 553;
Best Local Similarity 100.0%; Pred. No. 18;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 63 YSMIDTL 69
Db 458 YSMIDTL 464
RESULT 11
NIR_NEUCR
ID NIR_NEUCR STANDARD; PRT; 1176 AA.
AC P38681;
DT 01-FEB-1995 (Rel. 31, Created)
DT 01-FEB-1995 (Rel. 31, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Nitrite reductase [NAD(P)H] (EC 1.6.6.4).
GN NIT-6.
OS Neurospora crassa.
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
OC Sordariales; Sordariaceae; Neurospora.
OX NCBI_TaxID=5141;

```

```

RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-74-OR23-1A;
RX MEDLINE-93224461; PubMed-8096840;
RA Exley G.E., Colandene J.D., Garrett R.H.;
RT "Molecular cloning, characterization, and nucleotide sequence of
RT nit-6, the structural gene for nitrite reductase in Neurospora
RT crassa.";
RL J. Bacteriol. 175:2379-2392(1993).
CC -1- CATALYTIC ACTIVITY: 3 NAD(P)H + nitrite -> 3 NAD(P)(+) + NH(4)OH +
CC H(2)O.
CC -1- COFACTOR: THIS ENZYME IS A FAD FLAVOPROTEIN THAT ALSO CONTAINS
CC A SIROHEME AND ONE 2FE-2S IRON-SULFUR CENTER.
CC -1- PATHWAY: Nitrate assimilation (denitrification); second step.
CC -1- SUBUNIT: HOMODIMER (BY SIMILARITY).
CC -1- INDUCTION: BY NITRATE.
CC -1- SIMILARITY: THE C-TERMINAL DOMAIN IS A 4FE-4S/SIROHEME DOMAIN
CC FOUND IN NITRITE REDUCTASES (EC 1.6.6.4 AND EC 1.7.7.1) AND
CC SULFITE REDUCTASES (EC 1.8.1.2 AND EC 1.8.7.1).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; L07391; -; NOT_ANNOTATED_CDS.
DR PIR; A49848; A49848.
DR InterPro; IPR001327; FAD_Pyr_redox.
DR InterPro; IPR000660; Nit_sir.
DR InterPro; IPR005117; Nit_sir_fer.
DR InterPro; IPR001281; Rieske.
DR Pfam; PF00070; pyr_redox; 1.
DR Pfam; PF00355; Rieske; 1.
DR Pfam; PF01077; NIR_SIR; 1.
DR Pfam; PF03460; NIR_SIR_ferr; 1.
DR PRINTS; PR00397; SIROHAEM.
DR PRODOM; PD000139; FAD_pyr_redox; 1.
DR PROSITE; PS00365; NIR_SIR; 1.
KW Oxidoreductase; FAD; Flavoprotein; Iron-sulfur; Nitrate assimilation;
KW Heme; NADP.
FT NP_BIND 26 60 FAD (POTENTIAL).
FT NP_BIND 183 215 NAD(P)H (POTENTIAL).
FT METAL 717 717 IRON-SULFUR (2FE-2S) (POTENTIAL).
FT METAL 723 723 IRON-SULFUR (2FE-2S) (POTENTIAL).
FT METAL 757 757 IRON-SULFUR (2FE-2S) (POTENTIAL).
FT METAL 761 761 IRON-SULFUR (2FE-2S) AND SIROHEME
FT METAL (BY SIMILARITY).
FT DOMAIN 998 1054 PRO/SER-RICH.
FT SEQUENCE 1176 AA; 127367 MW; FFC7DCE66F80C710 CRC64;
SQ
Query Match 3.7%; Score 7; DB 1; Length 1176;
Best Local Similarity 100.0%; Pred. No. 35;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 135 IEKIKL 141
Db 41 IEKIKL 47
RESULT 12
PDRB_YEAST
ID PDRB_YEAST STANDARD; PRT; 1410 AA.
AC P40550; Q03092;
DT 01-FEB-1995 (Rel. 31, Created)
DT 01-FEB-1995 (Rel. 31, Last sequence update)
DT 01-OCT-1996 (Rel. 34, Last annotation update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE ATP-dependent permease PDR11.
GN PDR11 OR YIL013C.
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC

```

OC Saccharomycetales; Saccharomycetaceae; Saccharomyces.  
 OX NCBI\_TaxID=4932;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-S288c / AB972;  
 RA Barrall B.G., Badcock K., Bankier A.T., Bowman S., Brown D.,  
 RA Churcher C.M., Connor R., Copsey T., Dear S., Devlin K., Fraser A.,  
 RA Gentles S., Hamlyn N., Hornsall T.S., Hunt S., Jagels K., Jones M.,  
 RA Louis E., Lye G., Moule S., Moule T., Odell C., Pearson D.,  
 RA Rajandream M.A., Riles L., Rowley N., Skelton J., Smith V.,  
 RA Walsh S.V., Whitehead S.;  
 RL Submitted (DEC1994) to the EMBL/GenBank/DBJ databases.  
 RN [2]  
 RP IDENTIFICATION, AND SEQUENCE OF 1-10.  
 RX MEDLINE=95355421; PubMed=7629127;  
 RA Decottignies A., Lambert L., Catty P., Degand H., Eppling E.A.,  
 RA Moye-Rowley W.S., Balzi E., Goffeau A.;  
 RT "Identification and characterization of SNQ2, a new multidrug ATP  
 binding cassette transporter of the yeast plasma membrane.";  
 RL J. Biol. Chem. 270:18150-18157(1995).  
 CC -1 SUBCELLULAR LOCATION: Integral membrane protein (Potential).  
 CC -1 SIMILARITY: BELONGS TO THE ABC TRANSPORTER FAMILY. PDR5 SUBFAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL: Z38113; CAA86236.1; -  
 DR EMBL: Z46881; CAA86980.1; -  
 DR SGD: S0001275; PDR1.  
 DR InterPro: IPR003439; ABC\_transportr.  
 DR Pfam: PF00005; ABC\_tran; 2.  
 DR ProDom: PD00006; ABC\_transportr; 2.  
 DR PROSITE: PS00211; ABC\_TRANSPORTER; 1.  
 DR ATP-binding; Transmembrane; Glycoprotein; Transport.  
 KW INT\_MET 0  
 FT DOMAIN 1 387 0 CYTOPLASMIC (POTENTIAL).  
 FT TRANSMEM 388 408 POTENTIAL.  
 FT TRANSMEM 418 438 POTENTIAL.  
 FT TRANSMEM 471 491 POTENTIAL.  
 FT TRANSMEM 494 514 POTENTIAL.  
 FT TRANSMEM 524 544 POTENTIAL.  
 FT TRANSMEM 544 556 POTENTIAL.  
 FT TRANSMEM 556 636 POTENTIAL.  
 FT DOMAIN 636 1089 CYTOPLASMIC (POTENTIAL).  
 FT TRANSMEM 1090 1110 POTENTIAL.  
 FT TRANSMEM 1117 1137 POTENTIAL.  
 FT TRANSMEM 1175 1195 POTENTIAL.  
 FT TRANSMEM 1204 1224 POTENTIAL.  
 FT TRANSMEM 1230 1250 POTENTIAL.  
 FT TRANSMEM 1355 1375 POTENTIAL.  
 FT DOMAIN 1376 1410 CYTOPLASMIC (POTENTIAL).  
 FT NP\_BIND 781 788 ATP (POTENTIAL).  
 FT DOMAIN 707 712 POLY-SER.  
 FT DOMAIN 1045 1048 POLY-LEU.  
 FT CARBOHYD 594 594 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 1288 1288 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 1323 1323 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 1345 1345 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 SO SEQUENCE 1410 AA; 160405 MW; 93C9399A5CD14C3 CRC64;

Query Match 3.7%; Score 7; DB 1; Length 1410;  
 Best Local Similarity 100.0%; Pred. No. 42;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 19 FFIFLLE 25  
 DB 494 FFIFLLE 500

RESULT 13  
 YPRB\_SERMA STANDARD; PRT; 20 AA.  
 ID YPRB\_SERMA  
 AC P22581;  
 DT 01-AUG-1991 (Rel. 19, Created)  
 DT 01-AUG-1991 (Rel. 19, Last sequence update)  
 DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 DE Hypothetical protein in PROB 5'region (Fragment).  
 OS Serratia marcescens.  
 OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;  
 OC Serratia.  
 OX NCBI\_TaxID=615;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-Sr41;  
 RX MEDLINE=91237315; PubMed=1851803;  
 RA Omori K., Suzuki S., Inai Y., Komatsubara S.;  
 RT "Analysis of the Serratia marcescens proba operon and feedback  
 RT control of proline biosynthesis.";  
 RL J. Gen. Microbiol. 137:509-517(1991).  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL: D90351; BAA14363.1; -  
 DR EMBL: X53086; CAA37253.1; -  
 DR PIR: S11643; S11643.  
 DR PIR: C49753; C49753.  
 KW Hypothetical protein.  
 FT NON\_TER 1  
 FT SEQUENCE 20 AA; 2248 MW; 4DD7777735276674 CRC64;

Query Match 3.2%; Score 6; DB 1; Length 20;  
 Best Local Similarity 100.0%; Pred. No. 9.8;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 136 EKLKIL 141  
 DB 13 EKLKIL 18

RESULT 14  
 HSP\_CHICK STANDARD; PRT; 61 AA.  
 ID HSP\_CHICK  
 AC P15340; P02320;  
 DT 21-JUL-1986 (Rel. 01, Created)  
 DT 01-APR-1990 (Rel. 14, Last sequence update)  
 DT 15-JUN-2002 (Rel. 41, Last annotation update)  
 DE Sperm histone (Protamine) (Galline).  
 OS Gallus gallus (Chicken).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;  
 OC Gallus.  
 OX NCBI\_TaxID=9031;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=89308679; PubMed=2745455;  
 RA Oliva R., Dixon G.H.;  
 RT "Chicken protamine genes are intronless. The complete genomic  
 RT sequence and organization of the two loci.";  
 RL J. Biol. Chem. 264:12472-12481(1989).  
 RN [2]  
 RP SEQUENCE OF 48-61 FROM N.A.  
 RX MEDLINE=88112514; PubMed=2892748;  
 RA Oliva R., Mezquita J., Mezquita C., Dixon G.H.;  
 RT "Haploid expression of the rooster protamine mRNA in the postmeiotic  
 RT stages of spermatogenesis.";  
 RL Dev. Biol. 125:332-340(1988).

```

RN [3]
RP SEQUENCE.
RC TISSUE-Sperm;
RX MEDLINE=77050757; PubMed=992941;
RA Nakano M., Tobita T., Ando T.;
RT "Studies on a protamine (galline) from fowl sperm. 3. The total amino
   acid sequence of intact galline molecule."
RL Int. J. Pept. Protein Res. 8:565-578(1976).
CC -1- FUNCTION: PROTAMINES SUBSTITUTE FOR HISTONES IN THE CHROMATIN OF
   SPERM DURING THE HAPLOID PHASE OF SPERMATOGENESIS. THEY COMPACT
   SPERM DNA INTO A HIGHLY CONDENSED, STABLE AND INACTIVE COMPLEX.
CC -1- SUBCELLULAR LOCATION: Nuclear.
CC -1- TISSUE SPECIFICITY: TESTIS.
CC -1- SIMILARITY: BELONGS TO THE PROTAMINE P1 FAMILY.
CC -1- CAUTION: REF.3 SEQUENCE WAS VERY DIFFERENT FROM THAT OBTAINED BY
   REF.1 AND REF.2.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
   between the Swiss Institute of Bioinformatics and the EMBL outstation -
   the European Bioinformatics Institute. There are no restrictions on its
   use by non-profit institutions as long as its content is in no way
   modified and this statement is not removed. Usage by and for commercial
   entities requires a license agreement (See http://www.isb-sib.ch/announce/
   or send an email to license@isb-sib.ch).
CC -----
DR EMBL; L38713; AAA58721.1; -.
DR EMBL; M28100; AAA78951.1; -.
DR EMBL; M19078; AAA49049.1; ALT-SEQ.
DR PIR; A02662; GACH.
DR PIR; A34326; A34326.
DR InterPro; IPR000221; Protamine_P1.
DR Pfam; PF00260; Protamine_P1; 1.
DR PROSITE; PS00048; PROTAMINE_P1; 1.
KW Chromosomal protein; Nucleosome core; Spermatogenesis; DNA-binding;
FT Testis; DNA condensation; Nuclear protein; Phosphorylation.
FT INIT_MET 0 0 PHOSPHORYLATION (POTENTIAL).
FT MOD_RES 8 8
FT SEQUENCE 61 AA; 7986 MW; 191E27BD3A73AA3 CRC64;
SQ
Query Match 3.2%; Score 6; DB 1; Length 61;
Best Local Similarity 100.0%; Pred. No. 27;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 178 RSRRSV 183
Db 39 RSRRSV 44

```

```

RA Steiner J.M., Jakowitsch J., Bohnert H.J., Bryant D.A.;
RT "The complete sequence of the cyanelle genome of Cyanophora paradoxa:
   the genetic complexity of a primitive plastid."
RL (In) Schenk H.E.A., Herrmann R., Jeon K.W., Mueller N.E.,
   Schwemmler W. (eds.);
RL Eukaryotism and Symbiosis, pp.40-48, Springer-Verlag, Heidelberg
   (1997).
RL
CC -1- SIMILARITY: BELONGS TO THE YCF33 FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
   between the Swiss Institute of Bioinformatics and the EMBL outstation -
   the European Bioinformatics Institute. There are no restrictions on its
   use by non-profit institutions as long as its content is in no way
   modified and this statement is not removed. Usage by and for commercial
   entities requires a license agreement (See http://www.isb-sib.ch/announce/
   or send an email to license@isb-sib.ch).
CC -----
DR EMBL; U30821; AAA81285.1; -.
KW Cyanelle; Hypothetical protein.
SQ SEQUENCE 65 AA; 7635 MW; EAFDD0C8B6233CE3 CRC64;
QY 97 LIVLII 102
Db 40 LIVLII 45

```

```

RESULT 15
YCF33_CYAPA
ID YCF33_CYAPA STANDARD; PRT; 65 AA.
AC P48273;
DT 01-FEB-1996 (Rel. 33, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Hypothetical 7.6 kDa protein ycf33.
GN YCF33.
OS Cyanophora paradoxa.
OG Cyanelle.
OC Eukaryota; Glaucocystophyceae; Cyanophoraceae; Cyanophora.
OX NCBI_TaxID=2762;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=LB555 / Pringsheim;
RA Stirewalt V.L., Michalowski C.B., Loeffelhardt W., Bohnert H.J.,
   Bryant D.A.;
RT "Nucleotide sequence of the cyanelle DNA from Cyanophora paradoxa."
RL Plant Mol. Biol. Rep. 13:327-332(1995).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=LB555 / Pringsheim;
RA Loeffelhardt W., Stirewalt V.L., Michalowski C.B., Annarella M.,
   Farley J.Y., Schluchter W.M., Chung S., Newmann-Spallart C.,

```

```

RESULT 16
PSBH_PORPU
ID PSBH_PORPU STANDARD; PRT; 67 AA.
AC P51325;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 01-OCT-1996 (Rel. 34, Last annotation update)
DE Photosystem II 10 kDa phosphoprotein.
GN PSBH.
OS Porphyra purpurea.
OG Chloroplast.
OC Eukaryota; Rhodophyta; Bangiophyceae; Bangiales; Bangiaceae; Porphyra.
OX NCBI_TaxID=2787;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Avonport;
RA Reith M.E., Munholland J.;
RT "Complete nucleotide sequence of the Porphyra purpurea chloroplast
   genome."
RT Plant Mol. Biol. Rep. 13:333-335(1995).
CC -1- PTM: PHOSPHORYLATION IS A LIGHT DEPENDENT REACTION CATALYZED BY
   A MEMBRANE-BOUND KINASE (BY SIMILARITY).
CC -1- SIMILARITY: BELONGS TO THE PSBH FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
   between the Swiss Institute of Bioinformatics and the EMBL outstation -
   the European Bioinformatics Institute. There are no restrictions on its
   use by non-profit institutions as long as its content is in no way
   modified and this statement is not removed. Usage by and for commercial
   entities requires a license agreement (See http://www.isb-sib.ch/announce/
   or send an email to license@isb-sib.ch).
CC -----
DR EMBL; U38804; AAC08211.1; -.
DR InterPro; IPR001056; PSII_PsbH.
DR Pfam; PF00737; PsbH; 1.
DR ProDom; PD003584; PSII_PsbH; 1.
KW Photosystem II; Phosphorylation; Chloroplast; Transmembrane.
FT TRANSMEM 29 49 POTENTIAL.
FT SEQUENCE 67 AA; 7530 MW; 6CCB4C4AA40D3619 CRC64;
QY
Query Match 3.2%; Score 6; DB 1; Length 67;
Best Local Similarity 100.0%; Pred. No. 30;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

OY      151 SSVLE 156
      |||||
Db      51 SSVLE 56

RESULT 17
VG30_BPMD2
ID      VG30_BPMD2      STANDARD;      PRT;      88 AA.
AC      064224;
DT      15-DEC-1998 (Rel. 37, Created)
DT      15-DEC-1998 (Rel. 37, Last sequence update)
DE      15-DEC-1998 (Rel. 37, Last annotation update)
GN      Gene 30 protein (GP30).
OS      30.
OC      Mycobacteriophage D29.
OC      Viruses; dsDNA viruses, no RNA stage; Caudovirales; Siphoviridae;
OC      unclassified Siphoviridae.
OX      NCBI_TaxID=28369;
RN      [1]
RP      SEQUENCE FROM N.A.
RX      MEDLINE=98300335; PubMed=9636706;
RA      Ford M.E., Sarkis G.J., Belanger A.E., Hendrix R.W., Hatfull G.F.;
RT      "Genome structure of mycobacteriophage D29: implications for phage
RT      evolution.";
RL      J. Mol. Biol. 279:143-164(1998).
CC      -----
CC      This SWISS-PROT entry is copyright. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC      the European Bioinformatics Institute. There are no restrictions on its
CC      use by non-profit institutions as long as its content is in no way
CC      modified and this statement is not removed. Usage by and for commercial
CC      entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC      or send an email to license@isb-sib.ch).
CC      -----
DR      EMBL; AF022214; AAC18471.1;
SQ      SEQUENCE 88 AA; 10391 MW; 52071A2E676E5BCE CRC64;

Query Match      3.2%; Score 6; DB 1; Length 88;
Best Local Similarity 100.0%; Pred. No. 38;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY      117 IRLHE 122
      |||||
Db      65 IRLHE 70

RESULT 18
GAS4_ARATH
ID      GAS4_ARATH      STANDARD;      PRT;      106 AA.
AC      P46990; O49593;
DT      01-NOV-1995 (Rel. 32, Created)
DT      15-JUN-2002 (Rel. 41, Last sequence update)
DT      15-JUN-2002 (Rel. 41, Last annotation update)
DE      Gibberellin-regulated protein 4 precursor.
GN      GAS4 OR AT5G15230 OR F8M21_120.
OS      Arabidopsis thaliana (Mouse-ear cress).
OC      Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC      Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC      eustosids II; Brassicales; Brassicaceae; Arabidopsis.
OX      NCBI_TaxID=3702;
RN      [1]
RP      SEQUENCE FROM N.A.
RC      STRAIN=cv. C24; TISSUE=flower buds;
RX      MEDLINE=95244835; PubMed=7727751;
RA      Herzog M., Dorne A.-M., Grellet F.;
RT      "GAS4, a gibberellin-regulated gene family from Arabidopsis thaliana
RT      related to the tomato GAST1 gene.";
RL      Plant Mol. Biol. 27:743-752(1995).
RN      [2]
RP      SEQUENCE FROM N.A.
RX      MEDLINE=98179101; PubMed=9520278;
RA      Aubert D., Chevillard M., Dorne A.-M., Arlaud G., Herzog M.;

```

```

RT      "Expression patterns of GAS4 genes in Arabidopsis thaliana: the GAS4
RT      gene is up-regulated by gibberellins in meristematic regions.";
RL      Plant Mol. Biol. 36:871-883(1998).
RN      [3]
RP      SEQUENCE FROM N.A.
RC      STRAIN=cv. Columbia;
RX      MEDLINE=21016721; PubMed=11130714;
RA      Tabata S., Kaneko T., Nakamura Y., Kotani H., Kato T., Asanizu E.,
RA      Miyajima N., Sasamoto S., Kimura T., Hosouchi T., Kawashima K.,
RA      Kohara M., Matsumoto M., Muraki A., Nakayama S.,
RA      Nakazaki N., Naruo K., Okumura S., Shinpo S., Takeuchi C., Wada T.,
RA      Watanabe A., Yamada M., Yasuda M., Sato S., de la Bastide M.,
RA      Huang E., Spiegel L., Gnoj L., O'Shaughnessy A., Preston R.,
RA      Habermann K., Murray J., Johnson D., Rohlfing T., Nelson J.,
RA      Stoneking T., Pepin K., Spieth J., Sekhon M., Armstrong J., Becker M.,
RA      Belter E., Cordum H., Cordes M., Courtney L., Courtney W., Dante M.,
RA      Du H., Edwards J., Fryman J., Haakensen B., Lamar E., Latreille P.,
RA      Leonard S., Meyer R., Mulvaney E., Ozerky P., Riley A., Stromatt C.,
RA      Wagner-McPherson C., Wollam A., Yeakum M., Bell M., Dedhia N.,
RA      Parnell L., Shah R., Rodriguez M., Hoon See L., Vil D., Baker J.,
RA      Kirchhoff K., Toth K., King L., Bahret A., Miller B., Marra M.,
RA      Martienssen R., McCombie W.R., Wilson R.K., Murphy G., Bancroft I.,
RA      Volckaert G., Wambutt R., Duesterhoeft A., Stiekema W., Pohl T.,
RA      Entlan K.-D., Terryn N., Hartley N., Bent E., Johnson S.,
RA      Langham S.-A., McCullagh B., Robben J., Grymonprez B., Zimmermann W.,
RA      Ramsperger U., Wedler H., Balke K., Wedler E., Peters S.,
RA      van Staveren M., Dirkse W., Mooijman P., Klein Lankhorst R.,
RA      Welzenegger T., Bothe G., Rose M., Hauf J., Berneliser S., Hempel S.,
RA      Feldpausch M., Lamberth S., Villarroel R., Gietlen J., Ardiles W.,
RA      Bents O., Lemcke K., Kolesov G., Mayer K.F.X., Rudd S., Schoof H.,
RA      Scheller C., Zaccaria P., Mewes H.-W., Bevan M., Frnsz P.F.;
RT      "Sequence and analysis of chromosome 5 of the plant Arabidopsis
RT      thaliana.";
RL      Nature 408:823-826(2000).
RN      [4]
RP      SEQUENCE FROM N.A.
RC      STRAIN=cv. Columbia;
RA      Shinozaki K., Davis R.W., Ecker J.R., Theologis A.;
RT      "RIKEN Arabidopsis full length cDNA clones (RAFLs) sequenced by the
RT      SSP consortium (Salk/Stanford/PGECC).";
RL      Submitted (SEP-2001) to the EMBL/Genbank/DBJ databases.
CC      -1- SUBCELLULAR LOCATION: Secreted.
CC      -1- TISSUE SPECIFICITY: DEVELOPING ROOTS AND FLOWER BUDS.
CC      -1- INDUCTION: BY GIBBERELLINS.
CC      -1- PTM: SIX DISULFIDE BONDS MAY BE PRESENT.
CC      -1- SIMILARITY: BELONGS TO THE GAST1 FAMILY.
CC      -----
CC      This SWISS-PROT entry is copyright. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC      the European Bioinformatics Institute. There are no restrictions on its
CC      use by non-profit institutions as long as its content is in no way
CC      modified and this statement is not removed. Usage by and for commercial
CC      entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC      or send an email to license@isb-sib.ch).
CC      -----
DR      EMBL; U15683; AAA74480.1;
DR      EMBL; X98255; CAA66909.1;
DR      EMBL; AL353993; CAB89333.1;
DR      EMBL; AF360199; AAK25909.1;
DR      EMBL; AY057601; AAL14396.1;
DR      EMBL; AY040048; AAK64106.1;
DR      InterPro; IPR003854; GAS4.
DR      Pfam; PF02704; GAS4; 1.
KW      Multigene family; Signal.
FT      SIGNAL 1 25
FT      CHAIN 26 106
FT      CONFLICT 34 34 S -> R (IN REF. 1).
SQ      SEQUENCE 106 AA; 11996 MW; 63E7A1A964B07678 CRC64;

Query Match      3.2%; Score 6; DB 1; Length 106;
Best Local Similarity 100.0%; Pred. No. 46;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 95 LTLIVL 100
    |||||
Db 11 LTLIVL 16

RESULT 19
KV2D_MOUSE STANDARD; PRT; 112 AA.
ID KV2D_MOUSE
AC P01629;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE Ig kappa chain V-II region 2S1.3.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_Taxid=10090;
RN [1]
RP SEQUENCE.
RX MEDLINE=83055101; PubMed=7141411;
RA Herbst H., Chang J.Y., Abersold R., Braun D.G.;
RT "Murine VK25 isotype sequence: monoclonal antibody 2S1.3 specific for
RL Hoppe-Seyler's Z. Physiol. Chem. 363:1069-1076(1982).
CC -1- MISCELLANEOUS: THIS CHAIN IS FROM A HYBRIDOMA-DERIVED MONOCLONAL
CC ANTIBODY AGAINST THE STREPTOCOCCAL GROUP A POLYSACCHARIDE.
DR PIR; A01911; KVM51.
DR HSSP; P01607; 1RET.
DR InterPro; IPR003006; Ig_MHC.
DR Pfam; PF00047; Ig_V.
DR SMART; SM00406; IgV; 1.
KW Immunoglobulin V region; Monoclonal antibody.
FT DOMAIN 1 23 FRAMEWORK-1.
FT DOMAIN 24 39 COMPLEMENTARITY-DETERMINING-1.
FT DOMAIN 40 54 FRAMEWORK-2.
FT DOMAIN 55 61 COMPLEMENTARITY-DETERMINING-2.
FT DOMAIN 62 93 FRAMEWORK-3.
FT DOMAIN 94 102 COMPLEMENTARITY-DETERMINING-3.
FT DOMAIN 103 112 FRAMEWORK-4.
FT DISULFID 23 93 BY SIMILARITY.
FT NON_TER 112 112
SQ SEQUENCE 112 AA; 12221 MW; BD5FE56D789FBEC CRC64;

Query Match
Best Local Similarity 3.2%; Score 6; DB 1; Length 112;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 102 ITLYLV 107
    |||||
Db 35 ITLYLV 40

RESULT 20
OL7G_MOUSE STANDARD; PRT; 112 AA.
ID OL7G_MOUSE
AC Q60888;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 15-JUL-1998 (Rel. 36, Last annotation update)
DE Olfactory receptor 7G (M31) (Fragment).
GN OLFR7.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_Taxid=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=C57BL/6J;
RX MEDLINE=96149403; PubMed=8570653;
RA Sullivan S.L., Adamson M.C., Ressler K.J., Kozak C.A., Buck L.B.;
RT "The chromosomal distribution of mouse odorant receptor genes.";
RL Proc. Natl. Acad. Sci. U.S.A. 93:884-888(1996).

```

```

CC -1- FUNCTION: PUTATIVE ODORANT RECEPTOR.
CC -1- SUBCELLULAR LOCATION: Integral membrane protein.
CC -1- MISCELLANEOUS: OLFACTORY RECEPTORS 7A-7I ARE PRODUCED BY EIGHT
CC DIFFERENT GENES WITHIN THE OLFR7 COMPLEX.
CC -1- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; U28777; AAC52400.1; -.
DR MGD; MGI:104712; Olfr7.
DR InterPro; IPR000276; GPCR_Rhodpsn.
DR PROSITE; PS00237; G_PROTEIN_RECIP_F1_1; PARTIAL.
DR PROSITE; PS50262; G_PROTEIN_RECIP_F1_2; 1.
KW G-protein coupled receptor; Transmembrane; Multigene family;
KW Olfaction.
FT NON_TER 1 1
FT DOMAIN <1 12 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 13 33 4 (POTENTIAL).
FT DOMAIN 34 73 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 74 95 5 (POTENTIAL).
FT DOMAIN 96 109 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 110 >112 6 (POTENTIAL).
FT NON_TER 112 112
SQ SEQUENCE 112 AA; 12604 MW; 5C7E4942BA791F74 CRC64;

Query Match
Best Local Similarity 3.2%; Score 6; DB 1; Length 112;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 91 FFILTV 96
    |||||
Db 84 FFILTV 89

RESULT 21
KV2E_MOUSE STANDARD; PRT; 113 AA.
ID KV2E_MOUSE
AC P03976;
DT 23-OCT-1986 (Rel. 02, Created)
DT 23-OCT-1986 (Rel. 02, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE Ig kappa chain V-II region 17S29.1.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_Taxid=10090;
RN [1]
RP SEQUENCE.
RX TISSUE=Hybridoma;
RX MEDLINE=85128968; PubMed=6441768;
RA Abersold R., Herbst H., Grutter T., Chang J.Y., Braun D.G.;
RT "Murine V kappa 25 and V kappa 27 amino-acid sequences of C57BL/6
RT origin: monoclonal antibodies 17S29.1 and 22S25.1 specific for the
RT group A-streptococcal polysaccharide.";
RL Hoppe-Seyler's Z. Physiol. Chem. 365:1375-1383(1984).
CC -1- MISCELLANEOUS: ANTI-STREPTOCOCCAL GROUP A CARBOHYDRATE ANTIBODY.
CC PIR; A01912; KVM517.
DR HSSP; P01607; 1RET.
DR InterPro; IPR003006; Ig_MHC.
DR Pfam; PF00047; Ig_V.
DR SMART; SM00406; IgV; 1.
KW Immunoglobulin V region; Hybridoma.
FT DOMAIN 1 23 FRAMEWORK-1.
FT DOMAIN 24 39 COMPLEMENTARITY-DETERMINING-1.
FT DOMAIN 40 54 FRAMEWORK-2.
FT DOMAIN 55 61 COMPLEMENTARITY-DETERMINING-2.

```

```
FT DOMAIN 62 93 FRAMEWORK-3.
FT DOMAIN 94 102 COMPLEMENTARITY-DETERMINING-3.
FT DISULEID 103 112 FRAMEWORK-4.
FT NON_TER 23 93 BY SIMILARITY.
SQ SEQUENCE 113 AA; 12390 MW; 4E93797046F8DB33 CRC64;

Query Match
Best Local Similarity 100.0%; Score 6; DB 1; Length 113;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 102 ITLYLW 107
DB 35 ITLYLW 40

RESULT 22
RL30_SPOFR
ID RL30_SPOFR STANDARD; PRT; 113 AA.
AC P58375;
DT 15-JUN-2002 (Rel. 41, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DE 60S ribosomal protein L30.
GN RPL30.
OS Spodoptera frugiperda (Fall armyworm).
OC Eukaryota; Metazoa; Arthropoda; Mandibulata; Pancrustacea; Hexapoda;
OC Insecta; Pterygota; Neoptera; Endopterygota; Lepidoptera; Glossata;
OC Ditrysia; Noctuoidea; Noctuidae; Amphipyrinae; Spodoptera.
ON NCBI_TaxID=7108;
RN [1]
RP SEQUENCE FROM N.A.
RA Landais I., Ogliastro M., Mita K., Nohata J., Lopez-Ferber M.,
RA Duonor-Cerutti M., Fournier P., Devauchelle G.;
RT "Full-length ribosomal protein sequence from an EST library of
RT Spodoptera frugiperda cells (Sf9).";
RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: BELONGS TO THE L30E FAMILY OF RIBOSOMAL PROTEINS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: AF400193; AAK92165.1; -
CC InterPro: IPR000231; Ribosomal_L30e.
CC InterPro: IPR004038; Ribosomal_L7A.
CC Pfam: PF01248; Ribosomal_L7Ae; 1.
CC PROSITE: PS00709; RIBOSOMAL_L30E_1; 1.
CC PROSITE: PS00993; RIBOSOMAL_L30E_2; 1.
CC KW Ribosomal protein.
SQ SEQUENCE 113 AA; 12431 MW; B9A08AF6B8756AD8 CRC64;

Query Match
Best Local Similarity 100.0%; Score 6; DB 1; Length 113;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 34 CTLAIT 39
DB 92 CTLAIT 97

RESULT 23
RL30_BRABE
ID RL30_BRABE STANDARD; PRT; 114 AA.
AC P58374;
DT 15-JUN-2002 (Rel. 41, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE 60S ribosomal protein L30.
```

```
GN RPL30.
OS Branchiostoma belcheri (Amphioxus).
OC Eukaryota; Metazoa; Chordata; Cephalochordata; Branchiostomidae;
OC Branchiostoma.
ON NCBI_TaxID=7741;
RN [1]
RP SEQUENCE FROM N.A.
RA Chen Z., Zhang Y., Yang H., Zhang H., Han H., Li L., Wang X.;
RT "The primary structure of amphioxus ribosomal protein L30.";
RL Submitted (SEP-2001) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: BELONGS TO THE L30E FAMILY OF RIBOSOMAL PROTEINS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: AF420432; AAL09707.1; -
CC InterPro: IPR000231; Ribosomal_L30e.
CC InterPro: IPR004038; Ribosomal_L7A.
CC Pfam: PF01248; Ribosomal_L7Ae; 1.
CC PROSITE: PS00709; RIBOSOMAL_L30E_1; 1.
CC PROSITE: PS00993; RIBOSOMAL_L30E_2; 1.
CC KW Ribosomal protein.
SQ SEQUENCE 114 AA; 12689 MW; C4BFE184CE7C0096 CRC64;

Query Match
Best Local Similarity 100.0%; Score 6; DB 1; Length 114;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 34 CTLAIT 39
DB 89 CTLAIT 94

RESULT 24
ACPL_CASGL
ID ACPL_CASGL STANDARD; PRT; 136 AA.
AC P93092;
DT 15-JUN-2002 (Rel. 41, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Acyl carrier protein 1, chloroplast precursor (ACP 1).
GN ACP1.
OS Casuarina glauca (Swamp oak).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eucosids I; Fagales; Casuarinaceae; Casuarina.
ON NCBI_TaxID=3522;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Root nodules;
RA Laplaze L., Gherbi H., Franche C., Duhoux E., Bogusz D.;
RT "cDNA sequence for an acyl carrier protein from actinorhizal nodules
RT of Casuarina glauca.";
RL (In) Plant Gene Register PGR98-066.
CC -1- FUNCTION: THIS PROTEIN IS THE CARRIER OF THE GROWING FATTY ACID
CC CHAIN IN FATTY ACID BIOSYNTHESIS.
CC -1- SUBCELLULAR LOCATION: Chloroplast.
CC -1- SIMILARITY: CONTAINS 1 ACYL CARRIER DOMAIN.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: Y10994; CAA71885.1; -
```

DR HSSP; P02901; IACP.  
DR InterPro; IPR003231; Acyl\_carrier.  
DR InterPro; IPR003880; Ppantne\_attach.  
DR Pfam; PF00550; pp-binding; 1.  
DR ProDom; PD000887; Acyl\_carrier; 1.  
DR TIGRFAMs; TIGR00517; acyl\_carrier; 1.  
DR PROSITE; PS50075; ACP\_DOMAIN; 1.  
DR PROSITE; PS00012; PHOSPHOPANTHETHEINE; 1.  
DR Fatty acid biosynthesis; Phosphopantetheine; Chloroplast;  
KW Transit peptide; Multigene family.  
FT TRANSIT 1 52 CHLOROPLAST (POTENTIAL).  
FT CHAIN 53 136 ACYL CARRIER PROTEIN 1.  
FT BINDING 91 91 PHOSPHOPANTHETHEINE (BY SIMILARITY).  
SQ SEQUENCE 136 AA; 14351 MW; 8DF68F53079414FB CRC64;  
  
Query Match  
Best Local Similarity 3.2%; Score 6; DB 1; Length 136;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 134 LIEKLI 139  
|11111|  
Db 126 LIEKLI 131  
  
RESULT 25  
Y850\_AQUAE STANDARD; PRT; 138 AA.  
ID Y850\_AQUAE  
AC 067017;  
DT 16-OCT-2001 (Rel. 40, Created)  
DT 16-OCT-2001 (Rel. 40, Last sequence update)  
DT 16-OCT-2001 (Rel. 40, Last annotation update)  
DE Hypothetical protein AQ\_850.  
GN AQ\_850.  
OS Aquifex aeolicus.  
OC Bacteria; Aquificae; Aquificae (class); Aquificales; Aquificaceae;  
OC Aquifex.  
OX NCBI\_TaxID=63363;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=VF5;  
RX MEDLINE=98196666; PubMed=9537320;  
RA Deckert G., Warren P.V., Gaasterland T., Young W.G., Lenox A.L.,  
RA Graham D.E., Overbeek R., Sneed M.A., Keller M., Aujay M., Huber R.,  
RA Feldman R.A., Short J.M., Olson G.J., Swanson R.V.;  
RT "The complete genome of the hyperthermophilic bacterium Aquifex  
RT aeolicus.";  
RT Nature 392:353-358(1998).  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
CC  
CC EMBL; AE000710; AAC06980.1; -  
KW Hypothetical protein; Transmembrane; Complete proteome.  
FT TRANSMEM 11 33 POTENTIAL.  
SQ SEQUENCE 138 AA; 16457 MW; 7836DA14FC9C53F5 CRC64;  
  
Query Match  
Best Local Similarity 3.2%; Score 6; DB 1; Length 138;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 135 IEKLIK 140  
|11111|  
Db 77 IEKLIK 82  
  
RESULT 26  
Y4GB\_RHISN STANDARD; PRT; 139 AA.  
ID Y4GB\_RHISN

AC P55458;  
DT 01-NOV-1997 (Rel. 35, Created)  
DT 01-NOV-1997 (Rel. 35, Last sequence update)  
DT 01-NOV-1997 (Rel. 35, Last annotation update)  
DE Hypothetical 16.1 kDa protein Y4GB.  
GN Y4GB.  
OS Rhizobium sp. (strain NGR234).  
OG plasmid sym PNGR234a.  
OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;  
OC Rhizobiaceae; Rhizobium.  
OX NCBI\_TaxID=394;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=97305956; PubMed=9163424;  
RA Freiberg C.A., Fellay R., Bairoch A., Broughton W.J., Rosenthal A.,  
RA Perret X.;  
RT "Molecular basis of symbiosis between Rhizobium and legumes.";  
RL Nature 387:394-401(1997).  
CC -1- SIMILARITY: NONE OBVIOUS.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
CC  
CC EMBL; AE000074; AAB91676.1; -  
KW Hypothetical protein; Plasmid.  
SQ SEQUENCE 139 AA; 16113 MW; DBC300B06DC1F260 CRC64;  
  
Query Match  
Best Local Similarity 3.2%; Score 6; DB 1; Length 139;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 174 SLDLRS 179  
|11111|  
Db 19 SLDLRS 24  
  
RESULT 27  
ARI6\_CAEEL STANDARD; PRT; 146 AA.  
ID ARI6\_CAEEL  
AC P91167;  
DT 15-JUL-1998 (Rel. 36, Created)  
DT 15-JUL-1998 (Rel. 36, Last sequence update)  
DT 16-OCT-2001 (Rel. 40, Last annotation update)  
DE Probable ARP2/3 complex 16 kDa subunit (P16-ARC).  
GN C46H11.3.  
OS Caenorhabditis elegans.  
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditidae;  
OC Rhabditidae; Peloderinae; Caenorhabditis.  
OX NCBI\_TaxID=6239;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=Bristol N2;  
RA Miller N., Bradshaw H., Wamsley P.;  
RL Submitted (FEB-1997) to the EMBL/GenBank/DBJ databases.  
CC -1- FUNCTION: PART OF A COMPLEX IMPLICATED IN THE CONTROL OF ACTIN  
CC POLYMERIZATION IN CELLS (BY SIMILARITY).  
CC -1- SUBUNIT: BELONGS TO A COMPLEX COMPOSED OF ARP2, ARP3, P41-ARC,  
CC P34-ARC, P21-ARC, P20-ARC AND P16-ARC (BY SIMILARITY).  
CC -1- SIMILARITY: BELONGS TO THE ARPC5 FAMILY.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----

DR EMBL: U88314; AAF99882.1; -.  
DR WormRep; C46H11.3; CE08785.  
SQ SEQUENCE 146 AA; 16843 MW; F6897C272F10D77A CRC64;  
Query Match 3.2%; Score 6; DB 1; Length 146;  
Best Local Similarity 100.0%; Pred. No. 62;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 46 SADGP 51  
Db 26 SADGP 31  
RESULT 28  
SSB\_HUMAN STANDARD; PRT; 148 AA.  
AC Q04837;  
DT 01-OCT-1993 (Rel. 27, Created)  
DT 01-OCT-1993 (Rel. 27, Last sequence update)  
DT 15-JUN-2002 (Rel. 41, Last annotation update)  
DE Single-stranded DNA-binding protein, mitochondrial precursor  
(Mt-SSB) (MTSSB) (PWPI-interacting protein 17).  
GN SSBP1 OR SSBP.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=93246247; PubMed=8482537;  
RA Tiranti V., Rocchi M., Didonato S., Zeviani M.;  
RT Cloning of human and rat cDNAs encoding the mitochondrial single-  
stranded DNA-binding protein (SSB).";  
RL Gene 126:219-225(1993).  
RN [2]  
RP SEQUENCE FROM N.A.  
RA Honore B.;  
RT "hpw1-interacting protein 17 (ssDNA BP).";  
RL Submitted (JUN-2000) to the EMBL/GenBank/DBJ databases.  
RN [3]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Cervix;  
RA Strausberg R.;  
RL Submitted (NOV-2000) to the EMBL/GenBank/DBJ databases.  
RN [4]  
RX X-RAY CRYSTALLOGRAPHY (2.4 ANGSTROMS) OF 26-141.  
RA MEDLINE=97185916; PubMed=9033597;  
RA Yang C., Cuth U., Urbanke C., Kang C.;  
RT "Crystal structure of human mitochondrial single-stranded DNA binding  
protein at 2.4-A resolution.";  
RL Nat. Struct. Biol. 4:153-157(1997).  
CC -1- FUNCTION: THIS PROTEIN BINDS PREFERENTIALLY AND COOPERATIVELY  
TO SS-DNA. PROBABLY INVOLVED IN MITOCHONDRIAL DNA REPLICATION.  
CC -1- SUBUNIT: HOMOTETRAMER (BY SIMILARITY).  
CC -1- SUBCELLULAR LOCATION: Mitochondrial.  
CC -1- SIMILARITY: BELONGS TO THE SSB FAMILY.  
-----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
between the Swiss Institute of Bioinformatics and the EMBL outstation -  
the European Bioinformatics Institute. There are no restrictions on its  
use by non-profit institutions as long as its content is in no way  
modified and this statement is not removed. Usage by and for commercial  
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
-----  
CC EMBL: M94556; AAA6332.1; -.  
DR EMBL: AF277319; AAA69112.1; -.  
DR EMBL: BC000895; AAH00895.1; -.  
DR PIR: JN0568; JN0568.  
DR PDB: 3ULL; 15-OCT-97.  
DR Genew; HGNC:11317; SSBP1.  
DR MIM: 600439; -.  
DR InterPro; IPR00424; SSB\_protein.

DR Pfam: PF00436; SSB; 1.  
DR TIGRfams; TIGR00621; ssb; 1.  
DR PROSITE; PS00735; SSB\_1; 1.  
DR PROSITE; PS00736; SSB\_2; 1.  
KW DNA-binding; DNA replication; Mitochondrion; Transil peptide;  
KW 3D-structure.  
FT TRANSIT 1 16 MITOCHONDRION (BY SIMILARITY).  
FT CHAIN 17 148 SINGLE-STRANDED DNA-BINDING PROTEIN.  
SQ SEQUENCE 148 AA; 17260 MW; 98EE9E396D5636C2 CRC64;  
Query Match 3.2%; Score 6; DB 1; Length 148;  
Best Local Similarity 100.0%; Pred. No. 62;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 152 SLVLR 157  
Db 23 SLVLR 28  
RESULT 29  
Y919\_METJA STANDARD; PRT; 149 AA.  
ID Y919\_METJA  
AC Q58329;  
DT 15-JUL-1998 (Rel. 36, Created)  
DT 15-JUL-1998 (Rel. 36, Last sequence update)  
DT 16-OCT-2001 (Rel. 40, Last annotation update)  
DE Hypothetical protein MJ0919.  
GN MJ0919.  
OS Methanococcus jannaschii.  
OC Archaea; Euryarchaeota; Methanococci; Methanococcales;  
OC Methanocaldococcaceae; Methanocaldococcus.  
OX NCBI\_TaxID=2190;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=JAL-1 / DSM 2661 / ATCC 43067;  
RX MEDLINE=96337999; PubMed=8688087;  
RA Bult C.J., White O., Olsen G.J., Zhou L., Fleischmann R.D.,  
RA Sutton G.G., Blake J.A., Fitzgerald L.M., Clayton R.A., Gocayne J.D.,  
RA Kerlavage A.R., Dougherty B.A., Tomb J.-F., Adams M.D., Reich C.I.,  
RA Overbeek R., Kirkness E.F., Weinstock K.G., Merrick J.M., Glodek A.,  
RA Scott J.L., Geoghegan N.S.M., Weidman J.F., Fuhrmann J.L., Nguyen D.,  
RA Cotton M.D., Roberts K.M., Hurst M.A., Kaine B.P., Borodovsky M.,  
RA Uitterback T.R., Kelley J.M., Peterson J.D., Sadow P.W., Hanna M.C.,  
RA Klenk H.-P., Fraser C.M., Smith H.O., Woese C.R., Venter J.C.;  
RT "Complete genome sequence of the methanogenic archaeon, Methanococcus  
jannaschii.";  
RL Science 273:1058-1073(1996).  
-----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
between the Swiss Institute of Bioinformatics and the EMBL outstation -  
the European Bioinformatics Institute. There are no restrictions on its  
use by non-profit institutions as long as its content is in no way  
modified and this statement is not removed. Usage by and for commercial  
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
-----  
CC EMBL: U67535; AAB9927.1; -.  
DR DR TIGR: MJ0919; -.  
KW Hypothetical protein; Complete proteome.  
SQ SEQUENCE 149 AA; 16952 MW; 07B1A6483F957AE4 CRC64;  
Query Match 3.2%; Score 6; DB 1; Length 149;  
Best Local Similarity 100.0%; Pred. No. 63;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 135 IEKLIK 140  
Db 120 IEKLIK 125  
RESULT 30  
XY52\_PSEPU STANDARD; PRT; 157 AA.  
ID XY52\_PSEPU

```

AC Q05092;
DT 01-JUN-1994 (Rel. 29, Created)
DT 01-JUN-1994 (Rel. 29, Last sequence update)
DT 15-DEC-1998 (Rel. 37, Last annotation update)
DE XYLDLEGF operon transcriptional activator 2.
GN XYLS2.
OS Pseudomonas putida.
OG plasmid TOL PDK1, and plasmid TOL PWM53.
OC Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae;
OC Pseudomonas.
OX NCBI_TaxID=303;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=H51, and M53;
RX MEDLINE=93065221; PubMed=1331988;
RA Assinder S.J., de Marco P., Sayers J.R., Shaw L.E., Winson M.K.,
RA Williams P.A.;
RT "identical resolvases are encoded by Pseudomonas TOL plasmids PWM53
RT and PDK1."
RL Nucleic Acids Res. 20:5476-5476(1992).
CC -1- FUNCTION: REGULATORY PROTEIN OF THE TOL PLASMID XYL OPERONS. XYLS
CC ACTIVATES THE XYLYZLEGEFJOKH OPERON REQUIRED FOR THE DEGRADATION
CC OF TOLUENE, M-XYLENE AND P-XYLENE.
CC -1- SUBCELLULAR LOCATION: Cytoplasmic.
CC -1- SIMILARITY: BELONGS TO THE ARAC/XYLS FAMILY OF TRANSCRIPTIONAL
CC REGULATORS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; L02642; AAA71889.1; -
DR EMBL; L02643; AAA71891.1; -
DR PIR; S35486; S35486.
DR PIR; S35488; S35488.
DR InterPro; IPR000005; HTHARAC.
DR Pfam; PF00165; HTH_ARAC; 2.
DR PRINTS; PRO0032; HTHARAC.
DR SMART; SM00342; HTH_ARAC; 1.
DR PROSITE; PS00041; HTH_ARAC_FAMILY_1; 1.
DR PROSITE; PS01124; HTH_ARAC_FAMILY_2; 1.
KW Aromatic hydrocarbons catabolism; Transcription regulation; Activator;
KW DNA-binding; Plasmid.
FT DNA_BIND 55 74 H-T-H MOTIF (BY SIMILARITY).
SQ SEQUENCE 157 AA; 17894 MW; A5FB035F3F8F674 CRC64;
Query Match 3.2%; Score 6; DB 1; Length 157;
Best Local Similarity 100.0%; Pred. No. 66;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 164 GFLHLG 169
Db 115 GFLHLG 120

```

```

RP SEQUENCE FROM N.A.
RC TISSUE=Skeletal muscle;
RX MEDLINE=98010471; PubMed=9350988;
RA Valle G., Faulkner G., de Antoni A., Pacchioni B., Pallavicini A.,
RA Pandolfo D., Tiso N., Toppo S., Trevisan S., Lanfranchi G.;
RT "Telethonin, a novel sarcomeric protein of heart and skeletal
RT muscle."
RL FEBS Lett. 415:163-168(1997).
RN [2]
RP SEQUENCE FROM N.A.
RA Pallavicini A., Valle G., Lanfranchi G.;
RT "Human telethonin genomic sequence."
RL Submitted (AUG-1998) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RA Mues A., Gautel M.;
RT "Structure of the human telethonin gene."
RL Submitted (SEP-1998) to the EMBL/GenBank/DBJ databases.
RN [4]
RP SEQUENCE FROM N.A.
RC TISSUE=Prostate;
RA Strausberg R.;
RL Submitted (AUG-2001) to the EMBL/GenBank/DBJ databases.
RN [5]
RP INTERACTION WITH TITIN.
RX MEDLINE=98307394; PubMed=9645487;
RA Mues A., van der Ven P.F., Young P., Furst D.O., Gautel M.;
RT "Two immunoglobulin-like domains of the Z-disc portion of titin
RT interact in a conformation-dependent way with telethonin."
RL FEBS Lett. 428:111-114(1998).
CC -1- FUNCTION: MUSCLE ASSEMBLY REGULATING FACTOR.
CC -1- SUBUNIT: INTERACTS WITH TITIN.
CC -1- TISSUE SPECIFICITY: HEART AND SKELETAL MUSCLE.
CC -1- DISEASE: DEFECTS IN TCAP ARE A CAUSE OF LIMB-GIRDLE MUSCULAR
CC DYSTROPHY TYPE 2G (LGMD2G). TYPE 2 LIMB GIRDLE MUSCULAR
CC DYSTROPHIES REPRESENT A GENETICALLY HETEROGENEOUS GROUP OF
CC DISEASES WITH VARYING DEGREES OF SEVERITY DEPENDING ON AGE AT
CC ONSET AND RATE OF PROGRESSION. THE FEATURE THAT ALL THESE
CC CONDITIONS SHARE IS WEAKNESS STARTING IN THE PROXIMAL LIMB GIRDLE
CC MUSCULATURE.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AJ000491; CAA04129.1; -
DR EMBL; AJ010063; CAA08987.1; -
DR EMBL; AJ011098; CAA09479.1; -
DR EMBL; BC012628; AAH12628.1; -
DR Genew; HGNC:11610; TCAP.
DR MIM; 604488; -
DR MIM; 601954; -
SQ SEQUENCE 167 AA; 19052 MW; A3B0E27D8C84FC5 CRC64;
Query Match 3.2%; Score 6; DB 1; Length 167;
Best Local Similarity 100.0%; Pred. No. 70;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 68 TLSTRP 73
Db 29 TLSTRP 34

```

```

RESULT 31
TELT_HUMAN STANDARD; PRT; 167 AA.
ID TELT_HUMAN
AC 015273;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Telethonin (Titin cap protein).
GN TCAP.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN [1]

```

```

RESULT 32
TELT_MOUSE STANDARD; PRT; 167 AA.
ID TELT_MOUSE
AC 070548;
DT 30-MAY-2000 (Rel. 39, Created)

```

```

DT 30-MAY-2000 (Rel. 39, last sequence update)
DT 16-OCT-2001 (Rel. 40, last annotation update)
DE Telethonin (Titin cap protein).
GN TCAP.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Diaphragm;
RA Ievolella C., Formentin E., Valle G., Lanfranchi G.;
RT "Skeletal muscle transcripts characterization in Homo sapiens and Mus
RL musculus.",
RL Submitted (FEB-1998) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Heart;
RA Kolmerer B.;
RT "The titin cap protein - a novel protein essential for sarcomere
RT formation.",
RL Submitted (MAR-1999) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: MUSCLE ASSEMBLY REGULATING FACTOR.
CC -1- SUBUNIT: INTERACTS WITH TITIN.
CC -1- SUBCELLULAR LOCATION: SARCOMERIC.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: AJ233854; CA11585.1; -
DR EMBL: Y15845; CAB38077.1; -
DR MGD: MGI:1330233; Tcap.
SQ SEQUENCE 167 AA; 19078 MW; 2CB1F6F5415B4DC1 CRC64;

Query Match
Best Local Similarity 100.0%; Score 6; DB 1; Length 167;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 68 TLSTRP 73
Db 29 TLSTRP 34

RESULT 33
YC66_MESVI
ID YC66_MESVI STANDARD; PRT; 170 AA.
AC Q9M0Q8;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Hypothetical 19.4 kDa protein ycf66 (RF66).
GN YCF66.
OS Mesostigma viride.
OC Chloroplast.
OC Eukaryota; Viridiplantae; Streptophyta; Mesostigmatophyceae;
OC Mesostigmatales; Mesostigmataceae; Mesostigma.
OX NCBI_TaxID=41882;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=NIES-296;
RX MEDLINE=20150907; PubMed=10688199;
RA Lemieux C., Otis C., Turmel M.;
RT "Ancestral chloroplast genome in Mesostigma viride reveals an early
RT branch of green plant evolution.",
RL Nature 403:649-652(2000).
CC -1- SIMILARITY: BELONGS TO THE YCF66 FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration

```

```

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: AF166114; AAF43842.1; -
KW Chloroplast; Hypothetical protein.
SQ SEQUENCE 170 AA; 19397 MW; 912EA3682CC17459 CRC64;

Query Match
Best Local Similarity 100.0%; Score 6; DB 1; Length 170;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 FQPPSK 9
Db 110 FQPPSK 115

RESULT 34
RS5_TREPA
ID RS5_TREPA STANDARD; PRT; 172 AA.
AC O83236;
DT 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE 30S ribosomal protein S5.
GN RPSE OR TP0206.
OS Treponema pallidum.
OC Bacteria; Spirochaetales; Spirochaetaceae; Treponema.
OX NCBI_TaxID=160;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Nichols;
RX MEDLINE=98332770; PubMed=9665876;
RA Fraser C.M., Norris S.J., Weinstock G.M., White O., Sutton G.G.,
RA Dodson R., Gwin M., Hickey E.K., Clayton R., Ketchum K.A.,
RA Sodergren E., Hardham J.M., McLeod M.P., Salzberg S., Peterson J.,
RA Khalak H., Richardson D., Howell J.K., Chidambaram M., Uitterlinden T.,
RA McDonald L., Artach P., Bowman C., Cotton M.D., Fujii C., Garland S.,
RA Hatch B., Horst K., Roberts K., Sandusky M., Weidman J., Smith H.O.,
RA Venter J.C.;
RT "Complete genome sequence of Treponema pallidum, the syphilis
RT spirochete.",
RL Science 281:375-388(1998).
CC -1- FUNCTION: PROTEIN S5 IS IMPORTANT IN THE ASSEMBLY AND FUNCTION OF
CC THE 30S RIBOSOMAL SUBUNIT (BY SIMILARITY).
CC -1- SIMILARITY: BELONGS TO THE S5P FAMILY OF RIBOSOMAL PROTEINS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: AE001202; AAC65190.1; -
DR HSSP: P02357; 1PKP.
DR TIGR: TP0206; -
DR InterPro: IPR000851; Ribosomal_S5.
DR InterPro: IPR005324; Ribosomal_S5_C.
DR Pfam: PF00333; Ribosomal_S5_C; 1.
DR Pfam: PF03719; Ribosomal_S5_C; 1.
DR TIGRFAMS: TIGR01021; rpse_bact; 1.
DR PROSITE: PS00585; RIBOSOMAL_S5; 1.
KW Ribosomal protein; Complete proteome.
SQ SEQUENCE 172 AA; 18590 MW; FD8CB140449942A2 CRC64;

Query Match
Best Local Similarity 100.0%; Score 6; DB 1; Length 172;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 136 EKLK 141  
| | | | |  
Db 19 EKLK 24

## RESULT 35

```
SYC_BUCAP          STANDARD;          PRT;          177 AA.
ID SIB_BPP22
AC P38396;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Cysteine-1-trna synthetase (EC 6.1.1.16) (Cysteine--trna ligase)
DE (CysRS) (Fragment).
GN CYS.
OS Buchnera aphidicola (subsp. Schizaphis graminum).
OC Bacteria; Proteobacteria; gamma subdivision; Buchnera.
OX NCBI_TaxID=98794;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=95212914; PubMed=7535281;
RA Roubaksh D., Baumann P.;
RT "Characterization of a putative 23S-5S rRNA operon of Buchnera
RT aphidicola (endosymbiont of aphids) unlinked to the 16S rRNA-encoding
RT gene."
RL Gene 155:107-112(1995).
CC -1- CATALYTIC ACTIVITY: ATP + L-cysteine + trna(Cys) = AMP +
CC diphosphate + L-cysteine-1-trna(Cys).
CC -1- SUBUNIT: MONOMER (BY SIMILARITY).
CC -1- SUBCELLULAR LOCATION: Cytoplasmic.
CC -1- SIMILARITY: BELONGS TO CLASS-I AMINOACYL-TRNA SYNTHETASE FAMILY.
CC STRONG, TO METHIONYL-TRNA SYNTHETASE.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
-----
DR EMBL; U09230; AAD09434.1; -.
DR InterPro; IPR002308; Cys_trna-synt_1a.
DR InterPro; IPR001412; trna-synt_1.
DR Pfam; PF01406; trna-synt_1e; 1.
DR PROSITE; PS00178; AA_TRNA_LIGASE_I; FALSE_NEG.
KW Aminoacyl-trna synthetase; Protein biosynthesis; Ligase; ATP-binding.
FT SITE 29 40 "HIGH" REGION.
FT NON_TER 177 177
SQ SEQUENCE 177 AA; 20513 MW; D49CF11F1608EA19 CRC64;

Query Match 3.2%; Score 6; DB 1; Length 177;
Best Local Similarity 100.0%; Pred. No. 74;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

QY 135 IEKLIK 140  
| | | | |  
Db 125 IEKLIK 130

## RESULT 36

```
SIEB_BPP22          STANDARD;          PRT;          192 AA.
ID SIB_BPP22
AC P38396;
DT 01-OCT-1994 (Rel. 30, Created)
DT 01-OCT-1994 (Rel. 30, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Superinfection exclusion protein B.
GN SIEB OR GIT.
OS Bacteriophage P22.
OC Viruses; dsDNA viruses, no RNA stage; Caudovirales; Podoviridae;
OC P22-like viruses.
```

OX NCBI\_TaxID=10754;

RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=93328675; PubMed=8335629;  
RA Ranade K., Poteete A.R.;  
RT "Superinfection exclusion (sieb) genes of bacteriophages P22 and  
RT lambda."  
RL J. Bacteriol. 175:4712-4718(1993).

## RESULT 37

```
DHPS_CLOBE          STANDARD;          PRT;          205 AA.
ID DHPS_CLOBE
AC 005621;
DT 01-FEB-1995 (Rel. 31, Created)
DT 01-FEB-1995 (Rel. 31, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Dihydropteroate synthase (EC 2.5.1.15) (DHPS) (Dihydropteroate
DE pyrophosphorylase) (Fragment).
OS Clostridium beijerinckii (Clostridium MP).
OC Bacteria; Firmicutes; Clostridia; Clostridiales; Clostridiaceae;
OC Clostridium.
OX NCBI_TaxID=1520;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=ATCC 51743 / NCIMB 8052 / NCIB 8052;
RX MEDLINE=93380658; PubMed=8396545;
RA Oultiram J.D., Burr I.D., Elmore M.J., Minton N.P.;
RT "Cloning and sequence analysis of the genes encoding
RT phosphotransbutyrylase and butyrate kinase from Clostridium
RT acetobutylicum NCIMB 8052."
RL Gene 131:107-112(1993).
```

QY 97 LIVLII 102  
| | | | |  
Db 29 LIVLII 34

Query Match 3.2%; Score 6; DB 1; Length 192;  
Best Local Similarity 100.0%; Pred. No. 79;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
DR EMBL; L18800; AAA72007.1; -.
DR EMBL; AF217253; AAF75022.1; -.
DR EMBL; X02140; -. NOT_ANNOTATED_CDS.
DR PIR; A40606; A40606.
KW Transmembrane.
FT TRANSMEM 22 42 POTENTIAL.
FT TRANSMEM 51 71 POTENTIAL.
SQ SEQUENCE 192 AA; 22443 MW; D05756ED45649925 CRC64;

Query Match 3.2%; Score 6; DB 1; Length 192;
Best Local Similarity 100.0%; Pred. No. 79;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
-----
CC PHAGES THAT ARE INSENSITIVE TO REPRESSION.
CC -1- FUNCTION: HAS A ROLE IN THE PREVENTION OF SUPERINFECTION BY
CC -1- SUBCELLULAR LOCATION: Integral membrane protein (Probable).
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
-----
DR EMBL; L18800; AAA72007.1; -.
DR EMBL; AF217253; AAF75022.1; -.
DR EMBL; X02140; -. NOT_ANNOTATED_CDS.
DR PIR; A40606; A40606.
KW Transmembrane.
FT TRANSMEM 22 42 POTENTIAL.
FT TRANSMEM 51 71 POTENTIAL.
SQ SEQUENCE 192 AA; 22443 MW; D05756ED45649925 CRC64;
```

CC -1- FUNCTION: DHPS CATALYZES THE FORMATION OF THE IMMEDIATE PRECURSOR  
CC OF FOLIC ACID. IT IS IMPLICATED IN RESISTANCE TO SULFONAMIDE.  
CC -1- CATALYTIC ACTIVITY: 2-amino-4-hydroxy-6-hydroxymethyl-7,8-  
CC dihydropteridine diphosphate + 4-aminobenzoate = diphosphate +  
CC dihydropterate.  
CC -1- PATHWAY: Dihydrofolate biosynthesis; second step.  
CC -1- SIMILARITY: TO OTHER SPECIES DHPS.  
-----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
-----  
DR EMBL; L04468; AAA52083.1; -.  
DR PIR; P0620; P0620.  
DR HSP; 005701; 1AD1.  
DR InterPro; IPR000489; Dhdropt\_synt.  
DR Pfam; PF00809; Pterin\_bind; 1.  
DR PROSITE; PS00792; DHPS\_1; 1.  
DR PROSITE; PS00793; DHPS\_2; 1.  
KW Antibiotic resistance; Transferase; Folate biosynthesis.  
FT NON\_TER 205 205  
SQ SEQUENCE 205 AA; 22931 MW; 3609F76BC455ECBF CRC64;  
  
Query Match 3.2%; Score 6; DB 1; Length 205;  
Best Local Similarity 100.0%; Pred. NO. 84;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 70 STRPGY 75  
Db 60 STRPGY 65  
  
RESULT 38  
RAB4\_DICDI  
ID RAB4\_DICDI STANDARD; PRT; 205 AA.  
AC P36410;  
DT 01-JUN-1994 (Rel. 29, Created)  
DT 01-JUN-1994 (Rel. 29, Last sequence update)  
DT 01-NOV-1997 (Rel. 35, Last annotation update)  
DE Ras-related protein Rab4.  
GN RABD OR RAB4.  
OS Dictyostelium discoideum (Slime mold).  
OC Eukaryota; Mycetozoa; Dictyostelida; Dictyostelium.  
OX NCBI\_TaxID=44689;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=AX3;  
RX MEDLINE=95181582; PubMed=7876348;  
RA Bush J.M. IV, Nolta K., Rodriguez-Paris J., Kaufmann N.,  
RA O'Halloran T., Ruscetti T., Temesvari L., Steck T., Cardelli J.A.;  
RT "A Rab4-like GTPase in Dictyostelium discoideum colocalizes with  
RT V-H(+)-ATPases in reticular membranes of the contractile vacuole  
RT complex and in lysosomes."  
RL J. Cell Sci. 107:2801-2812(1994).  
CC -1- SUBCELLULAR LOCATION: ASSOCIATED WITH THE SPONGIOMES OF  
CC CONTRACTILE VACUOLE COMPLEX.  
CC -1- SIMILARITY: TO RAS PROTEINS. BELONGS TO THE RAB SUBFAMILY.  
-----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
-----  
DR EMBL; U02927; AAA80151.1; -.  
DR HSP; P21181; 1AM4.  
DR Dictydb; DD00069; rabd.

DR InterPro; IPR003579; GTPase\_Rab.  
DR InterPro; IPR001806; Ras\_trnsfmng.  
DR InterPro; IPR005225; Small\_GTP.  
DR Pfam; PF00071; ras; 1.  
DR PRINTS; PR00449; RASTRNSFRMNG.  
DR SMART; SM00175; RAB; 1.  
DR TIGRfams; TIGR00231; small\_gtp; 1.  
KW GTP-binding; Lipoprotein; Prenylation.  
FT NP\_BIND 15 22 GTP (BY SIMILARITY).  
FT NP\_BIND 63 67 GTP (BY SIMILARITY).  
FT NP\_BIND 121 124 GTP (BY SIMILARITY).  
FT DOMAIN 37 45 EFFECTOR REGION (BY SIMILARITY).  
FT LIPID 204 204 GERANYL-GERANYL (BY SIMILARITY).  
SQ SEQUENCE 205 AA; 23192 MW; 2BF51383A73724F1 CRC64;  
  
Query Match 3.2%; Score 6; DB 1; Length 205;  
Best Local Similarity 100.0%; Pred. NO. 84;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 182 SVOEGN 187  
Db 172 SVOEGN 177  
  
RESULT 39  
KGUA\_THEMA  
ID KGUA\_THEMA STANDARD; PRT; 207 AA.  
AC Q9X215;  
DT 16-OCT-2001 (Rel. 40, Created)  
DT 16-OCT-2001 (Rel. 40, Last sequence update)  
DT 16-OCT-2001 (Rel. 40, Last annotation update)  
DE Guanylate kinase (EC 2.7.4.8) (GMP kinase).  
GN GMK OR TM1689.  
OS Thermotoga maritima.  
OC Bacteria; Thermotogae; Thermotogales;  
OC Thermotogaceae; Thermotoga.  
OX NCBI\_TaxID=2336;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=MSB8 / DSM 3109;  
RX MEDLINE=99287316; PubMed=10360571;  
RA Nelson K.E., Clayton R.A., Gill S.R., Gwin M.L., Dodson R.J.,  
RA Haft D.H., Hickey E.K., Peterson J.D., Nelson W.C., Ketchum K.A.,  
RA McDonald L., Utterback T.R., Malek J.A., Linher K.D., Garrett M.M.,  
RA Stewart A.M., Cotton M.D., Pratt M.S., Phillips C.A., Richardson D.,  
RA Heidelberg J., Sutton G.G., Fleischmann R.D., Eisen J.A., White O.,  
RA Salzberg S.L., Smith H.O., Venter J.C., Fraser C.M.;  
RT "Evidence for lateral gene transfer between Archaea and Bacteria from  
RT genome sequence of Thermotoga maritima."  
RL Nature 399:323-329(1999).  
CC -1- FUNCTION: ESSENTIAL FOR RECYCLING GMP AND INDIRECTLY, CGMP.  
CC -1- CATALYTIC ACTIVITY: ATP + GMP = ADP + GDP.  
CC -1- SUBCELLULAR LOCATION: CYTOPLASMIC (BY SIMILARITY).  
CC -1- SIMILARITY: BELONGS TO THE GUANYLATE KINASE FAMILY.  
-----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
-----  
DR EMBL; AE001809; AAD36756.1; -.  
DR HSP; P15454; 1GKY.  
DR TIGR; TM1689; -.  
DR InterPro; IPR000619; Guanylate\_kin.  
DR Pfam; PF00625; Guanylate\_kin; 1.  
DR SMART; SM00072; Gukc; 1.  
DR PROSITE; PS00856; GUANYLATE\_KINASE\_1; 1.  
DR PROSITE; PS50052; GUANYLATE\_KINASE\_2; 1.  
KW Transferase; Kinase; ATP-binding; Complete proteome.  
FT NP\_BIND 10 17 ATP (BY SIMILARITY).

SEQ SEQUENCE 207 AA; 24037 MW; E38375CFE2E26963 CRC64;  
Query Match 3.2%; Score 6; DB 1; Length 207;  
Best Local Similarity 100.0%; Pred. No. 85;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 125 INEGKD 130  
Db 90 INEGKD 95  
RESULT 40  
RHOL\_ENTHI STANDARD; PRT; 208 AA.  
ID RHOL\_ENTHI  
AC P31021;  
DT 01-JUL-1993 (Rel. 26, Created)  
DT 01-JUL-1993 (Rel. 26, Last sequence update)  
DT 01-NOV-1995 (Rel. 32, Last annotation update)  
DE RAS-like GTP-binding protein RHOL.  
GN RHOL.  
OS Entamoeba histolytica.  
OC Eukaryota; Entamoebidae; Entamoeba.  
OX NCBI\_TaxID=5759;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX STRAIN=HM-1:IMSS;  
RX MEDLINE=93211449; PubMed=8459832;  
RA Lohia A., Samuelson J.;  
RT "Molecular cloning of a rho family gene of Entamoeba histolytica";  
RL Mol. Biochem. Parasitol. 58:177-180(1993).  
CC -1- SIMILARITY: BELONGS TO THE SMALL GTPASE SUPERFAMILY. RHO FAMILY.  
CC -1- CAUTION: IT IS UNCERTAIN WHETHER MET-1 OR MET-7 IS THE INITIATOR.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL; L03809; AAA29114.1; -;  
DR EMBL; L03809; AAA29115.1; -;  
DR HSSP; P21181; IMA4.  
DR DR InterPro; IPR003578; GTPase\_Rho.  
DR DR InterPro; IPR001806; Ras\_trnsfrmng.  
DR DR InterPro; IPR005225; Small\_GTP.  
DR Pfam; PF00071; Ras; 1.  
DR SMART; SM00174; RHO; 1.  
DR DR TIGRFAMS; TIGR00231; small\_GTP; 1.  
KW GTP-binding.  
FT NP\_BIND 27 34 GTP (BY SIMILARITY).  
FT NP\_BIND 74 78 GTP (BY SIMILARITY).  
FT NP\_BIND 132 135 GTP (BY SIMILARITY).  
FT DOMAIN 49 57 EFFECTOR REGION (POTENTIAL).  
SQ SEQUENCE 208 AA; 23220 MW; AF24D5821717200F CRC64;  
Query Match 3.2%; Score 6; DB 1; Length 208;  
Best Local Similarity 100.0%; Pred. No. 85;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 19 FFIFLL 24  
Db 201 FFIFLL 206  
RESULT 41  
CCMB\_PARDE STANDARD; PRT; 215 AA.  
ID CCMB\_PARDE  
AC P52219;  
DT 01-OCT-1996 (Rel. 34, Created)  
DT 01-OCT-1996 (Rel. 34, Last sequence update)  
DT 15-JUN-2002 (Rel. 41, Last annotation update)

DE Heme exporter protein B (Cytochrome c-type biogenesis protein ccmB).  
GN CCMB.  
OS Paracoccus denitrificans.  
OC Bacteria; Proteobacteria; alpha subdivision; Rhodobacter group;  
OC Paracoccus.  
OX NCBI\_TaxID=266;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX STRAIN=Pd 1222;  
RX MEDLINE=97195802; PubMed=9043133;  
RA Page D., Pearce D.A., Norris H.A., Ferguson S.J.;  
RT "The Paracoccus denitrificans ccmA, B and C genes: cloning and  
RT sequencing, and analysis of the potential of their products to form a  
RT haem or apo-c-type cytochrome transporter."  
RL Microbiology 143:563-576(1997).  
CC -1- FUNCTION: REQUIRED FOR THE EXPORT OF HEME TO THE PERIPLASM FOR THE  
CC BIOGENESIS OF C-TYPE CYTOCHROMES (BY SIMILARITY).  
CC -1- SUBCELLULAR LOCATION: Integral membrane protein. Inner membrane  
CC (Probable).  
CC -----  
CC -1- SIMILARITY: BELONGS TO THE CCMB/CYCW/HELB FAMILY.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL; Z71971; CAA96494.1; -;  
DR InterPro; IPR003544; Cyt\_c\_biog\_CcMB.  
DR Pfam; PF03379; CcMB; 1.  
DR PRINTS; PR01414; CCMBBIOGENSIS.  
DR TIGRFAMS; TIGR01190; ccmB; 1.  
KW Cytochrome c-type biogenesis; Transport; Transmembrane;  
KW Inner membrane.  
FT TRANSMEM 17 37 POTENTIAL.  
FT TRANSMEM 47 67 POTENTIAL.  
FT TRANSMEM 98 118 POTENTIAL.  
FT TRANSMEM 125 145 POTENTIAL.  
FT TRANSMEM 154 174 POTENTIAL.  
FT TRANSMEM 190 210 POTENTIAL.  
SQ SEQUENCE 215 AA; 21960 MW; 9B882459B41E1D91 CRC64;  
Query Match 3.2%; Score 6; DB 1; Length 215;  
Best Local Similarity 100.0%; Pred. No. 88;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 172 DGSIDL 177  
Db 73 DGSIDL 78  
RESULT 42  
RB14\_HUMAN STANDARD; PRT; 215 AA.  
ID RB14\_HUMAN  
AC P35287; Q9U111; Q969L0;  
DT 01-FEB-1994 (Rel. 28, Created)  
DT 01-FEB-1994 (Rel. 28, Last sequence update)  
DT 15-JUN-2002 (Rel. 41, Last annotation update)  
DE Ras-related protein Rab-14.  
GN RAB14.  
OS Homo sapiens (Human), and  
OS Rattus norvegicus (Rat).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
OX NCBI\_TaxID=9606, 10116;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC SPECIES=Human; TISSUE=Liver;  
RA Proikas-Cezanne T., Jenkins J.R.;  
RT "Human Rab14 cloning and intracellular localization to the  
RT biosynthetic/secretory trafficking pathway.";

RL Submitted (MAY-1999) to the EMBL/GenBank/DBJ databases.  
RN [2]  
RP SEQUENCE FROM N.A.  
RC SPECIES=Human;  
RA Ren Y.;  
RT "Cloning and characterization of human small GTPase Rab14.";  
RL Submitted (NOV-1999) to the EMBL/GenBank/DBJ databases.  
RN [3]  
RP SEQUENCE FROM N.A.  
RC SPECIES=Human; TISSUE=adrenal gland;  
RX MEDLINE=20402571; PubMed=10931946;  
RA Hu R.-M., Han Z.-G., Song H.-D., Peng Y.-D., Huang Q.-H., Ren S.-X.,  
RA Gu Y.-J., Huang C.-H., Li Y.-B., Jiang C.-L., Fu G., Zhang Q.-H.,  
RA Gu B.-W., Dai M., Mao Y.-F., Gao G.-F., Rong R., Ye M., Zhou J.,  
RA Xu S.-H., Gu J., Shi J.-X., Jin W.-R., Zhang C.-K., Wu T.-M.,  
RA Huang G.-Y., Chen Z., Chen M.-D., Chen J.-L.;  
RT "Gene expression profiling in the human hypothalamus-pituitary-adrenal  
RT axis and full-length cDNA cloning.";  
RL Proc. Natl. Acad. Sci. U.S.A. 97:9543-9548(2000).  
RN [4]  
RP SEQUENCE FROM N.A.  
RC SPECIES=Human; TISSUE=Melanoma;  
RA Blum H., Bauersachs S., Mewes H.-W., Weil B., Wiemann S.;  
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.  
RN [5]  
RP SEQUENCE FROM N.A.  
RC SPECIES=Human; TISSUE=Placenta;  
RA Isogai T., Ota T., Hayashi K., Sugiyama T., Otsuki T., Suzuki Y.,  
RA Nishikawa T., Nagai K., Sugano S., Takahashi-Fujii A., Hara H.,  
RA Tanase T., Nomura Y., Togiya S., Komai F., Hara R., Takeuchi K.,  
RA Arita M., Nabekura T., Ishii S., Kawai Y., Saito K., Yamamoto J.,  
RA Wakamatsu A., Nakamura Y., Nagahari K., Masuno Y., Oshima A.;  
RT "NEDO human cDNA sequencing project.";  
RL Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.  
RN [6]  
RP SEQUENCE FROM N.A.  
RC SPECIES=Human;  
RA Laird G.;  
RL Submitted (JAN-2002) to the EMBL/GenBank/DBJ databases.  
RN [7]  
RP SEQUENCE FROM N.A.  
RC SPECIES=Human; TISSUE=Placenta;  
RA Strausberg R.;  
RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.  
RN [8]  
RP SEQUENCE FROM N.A.  
RC SPECIES=Rat; STRAIN=Sprague-Dawley; TISSUE=Brain;  
RX MEDLINE=92210533; PubMed=1313420;  
RA Elferink L.A., Anzai K., Scheller R.H.;  
RT "Rab15, a novel low molecular weight GTP-binding protein specifically  
RT expressed in rat brain.";  
RL J. Biol. Chem. 267:5768-5775(1992).  
CC -1- FUNCTION: MAY BE INVOLVED IN VESICULAR TRAFFICKING AND  
CC NEUROTRANSMITTER RELEASE.  
CC -1- TISSUE SPECIFICITY: HIGHEST LEVELS FOUND IN WHOLE BRAIN, SPINAL  
CC CORD, HEART, KIDNEY AND LUNG.  
CC -1- SIMILARITY: TO RAS PROTEINS. BELONGS TO THE RAB SUBFAMILY.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL; AF152463; AAF00150.1; -  
DR EMBL; AF203689; AAF19400.1; -  
DR EMBL; AF112206; AAF17194.1; -  
DR EMBL; AL162081; CAB82414.1; -  
DR EMBL; AK023524; BAB14598.1; -  
DR EMBL; AL137068; CAD20124.1; -  
DR EMBL; BC006081; AAH06081.1; -

DR EMBL; M83680; AAA41994.1; -  
DR PIR; E42148; E42148.  
DR HSSP; P36017; 1EKO.  
DR Genew; HGNC:16524; RAB14.  
DR InterPro; IPR003579; GTPase\_Rab.  
DR InterPro; IPR001806; Ras\_trnsfrmng.  
DR InterPro; IPR005225; Small\_GTP.  
DR Pfam; PF00071; ras; 1.  
DR PRINTS; PRO0449; RASTRNSFRMG.  
DR SMART; SM00175; RAB; 1.  
DR TIGRFRAMS; TIGR00231; small\_GTP; 1.  
KW GTP-binding; Lipoprotein; Prenylation; Protein transport;  
KW Polymorphism.  
FT NP\_BIND 18 25 GTP (BY SIMILARITY).  
FT NP\_BIND 66 70 GTP (BY SIMILARITY).  
FT NP\_BIND 124 127 GTP (BY SIMILARITY).  
FT DOMAIN 40 48 EFFECTOR REGION (BY SIMILARITY).  
FT LIPID 213 213 GERANYL-GERANYL (BY SIMILARITY).  
FT LIPID 215 215 GERANYL-GERANYL (BY SIMILARITY).  
FT LIPID VARIANT 4 4 T -> A.  
SQ SEQUENCE 215 AA; 23927 MW; B435AAC6F562B007 CRC64;  
/FTID=VAR\_012986.  
Query Match 3.2%; Score 6; DB 1; Length 215;  
Best Local Similarity 100.0%; Pred. No. 88;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 172 DGLDL 177  
Db 178 DGLDL 183  
RESULT 43  
CCMB\_RHOCA ID CCMB\_RHOCA STANDARD; PRT; 218 AA.  
AC P29960;  
DT 01-APR-1993 (Rel. 25, Created)  
DT 01-APR-1993 (Rel. 25, Last sequence update)  
DT 15-JUN-2002 (Rel. 41, Last annotation update)  
DE Heme exporter protein B (Cytochrome c-type biogenesis protein helb).  
GN HELB.  
OS Rhodobacter capsulatus (Rhodospirillum rubrum).  
OC Bacteria; Proteobacteria; alpha subdivision; Rhodobacter group;  
OC Rhodobacter.  
OX NCBI\_TaxID=1061;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=SB1003 / St Louis;  
RX MEDLINE=92146961; PubMed=1310666;  
RA Beckman D.L., Trawick D.R., Kranz R.G.;  
RT "Bacterial cytochromes c biogenesis.";  
RL Genes Dev. 6:268-283(1992).  
CC -1- FUNCTION: REQUIRED FOR THE EXPORT OF HEME TO THE PERIPLASM FOR THE  
CC BIOGENESIS OF C-TYPE CYTOCHROMES.  
CC -1- SUBCELLULAR LOCATION: Integral membrane protein. Inner membrane  
CC (Probable).  
CC -1- SIMILARITY: BELONGS TO THE CCMB/CYCW/HELB FAMILY.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL; X63462; CAA45062.1; -  
DR PIR; S23664; S23664.  
DR InterPro; IPR003544; Cytc\_biolg\_CcMB.  
DR Pfam; PF03379; CcMB; 1.  
DR PRINTS; PRO1414; CCMBIOGNIS.  
DR TIGRFRAMS; TIGR01190; ccmb; 1.  
KW Cytochrome c-type biogenesis; Transport; Transmembrane;

```
KW Inner membrane.
FT TRANSMEM 18 38 POTENTIAL.
FT TRANSMEM 43 63 POTENTIAL.
FT TRANSMEM 99 119 POTENTIAL.
FT TRANSMEM 131 151 POTENTIAL.
FT TRANSMEM 155 175 POTENTIAL.
FT TRANSMEM 185 205 POTENTIAL.
SQ SEQUENCE 218 AA; 22611 MW; CA2382FBD4B447F4 CRC64;

Query Match
Best Local Similarity 3.2%; Score 6; DB 1; Length 218;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 172 DGSIDL 177
Db 74 DGSIDL 79

RESULT 44
Y008_METJA STANDARD; PRT; 220 AA.
ID Y008_METJA STANDARD; PRT; 220 AA.
AC 060319;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Hypothetical protein MJ0008.
GN MJ0008.
OS Methanococcus jannaschii.
OC Archaea; Euryarchaeota; Methanococci; Methanococcales;
OC Methanocaldococcaceae; Methanocaldococcus.
OX NCBI_TaxID=2190;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=JAL-1 / DSM 2661 / ATCC 43067;
RX MEDLINE=96337999; PubMed=8688087;
RA Bult C.J., White O., Olsen G.J., Zhou L., Fleischmann R.D.,
RA Sutton G.G., Blake J.A., Fitzgerald L.M., Clayton R.A., Gocayne J.D.,
RA Kerlavage A.R., Dougherty B.A., Tomb J.-F., Adams M.D., Reich C.I.,
RA Overbeek R., Kirkness E.F., Weinstock K.G., Merrick J.M., Glodek A.,
RA Scott J.L., Geoghagen N.S.M., Weidman J.F., Fuhrmann J.L., Nguyen D.,
RA Uitterback T.R., Kelley J.M., Peterson J.D., Sadow P.W., Hanna M.C.,
RA Cotton M.D., Roberts K.M., Hurst M.A., Kaine B.P., Borodovsky M.,
RA Klenk H.-P., Fraser C.M., Smith H.O., Woese C.R., Venter J.C.;
RT "Complete genome sequence of the methanogenic archaeon, Methanococcus
jannaschii."
RL Science 273:1058-1073(1996).
CC -1- SIMILARITY: SOME, TO M.JANNASCHII MJ1311.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; U67460; AAB97997.1; -.
CC DR TIGR; MJ0008; -.
CC KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 220 AA; 25836 MW; 955DF4A0FAB71B88 CRC64;

Query Match
Best Local Similarity 3.2%; Score 6; DB 1; Length 220;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 131 KMFLEI 136
Db 193 KMFLEI 198

RESULT 45
HYPB_METJA STANDARD; PRT; 221 AA.
ID HYPB_METJA STANDARD; PRT; 221 AA.
```

```
AC Q57884;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Probable hydrogenase nickel incorporation protein hypb.
GN HYPB OR MJ0442.
OS Methanococcus jannaschii.
OC Archaea; Euryarchaeota; Methanococci; Methanococcales;
OC Methanocaldococcaceae; Methanocaldococcus.
OX NCBI_TaxID=2190;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=JAL-1 / DSM 2661 / ATCC 43067;
RX MEDLINE=96337999; PubMed=8688087;
RA Bult C.J., White O., Olsen G.J., Zhou L., Fleischmann R.D.,
RA Sutton G.G., Blake J.A., Fitzgerald L.M., Clayton R.A., Gocayne J.D.,
RA Kerlavage A.R., Dougherty B.A., Tomb J.-F., Adams M.D., Reich C.I.,
RA Overbeek R., Kirkness E.F., Weinstock K.G., Merrick J.M., Glodek A.,
RA Scott J.L., Geoghagen N.S.M., Weidman J.F., Fuhrmann J.L., Nguyen D.,
RA Uitterback T.R., Kelley J.M., Peterson J.D., Sadow P.W., Hanna M.C.,
RA Cotton M.D., Roberts K.M., Hurst M.A., Kaine B.P., Borodovsky M.,
RA Klenk H.-P., Fraser C.M., Smith H.O., Woese C.R., Venter J.C.;
RT "Complete genome sequence of the methanogenic archaeon, Methanococcus
jannaschii."
RL Science 273:1058-1073(1996).
CC -1- FUNCTION: COULD BE INVOLVED IN NICKEL BINDING AND ACCUMULATION.
CC -1- SIMILARITY: BELONGS TO THE HYPB/HUPM FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; U67495; AAB98429.1; -.
CC DR TIGR; MJ0442; -.
CC DR InterPro; IPR004392; HYPB.
CC DR InterPro; IPR002894; HYPB_UREG.
CC DR Pfam; PF01495; HYPB_UREG; 1.
CC DR TIGRFAMS; TIGR00073; hypb; 1.
CC KW Metal-binding; Nickel; Complete proteome.
SQ SEQUENCE 221 AA; 24340 MW; 0BB5415386A81138 CRC64;

Query Match
Best Local Similarity 3.2%; Score 6; DB 1; Length 221;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 134 LIEKLI 139
Db 49 LIEKLI 54
```

Search completed: November 9, 2002, 07:28:38  
Job time : 41 secs